FIRST- PRINCIPLES MOLECULAR-DYNAMICS STUDY OF METAL-SUPPORTED NANOSYSTEMS

THÈSE N^{O} 3026 (2004)

PRÉSENTÉE À LA FACULTÉ SCIENCES DE BASE

Institut de théorie des phénomènes physiques

SECTION DE PHYSIQUE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Massimiliano STENGEL

laurea in fisica, Università degli Studi di Trieste, Italie
et de nationalité italienne

acceptée sur proposition du jury:

Prof. A. Baldereschi, directeur de thèse
Dr W. Andreoni, rapporteur
Dr A. De Vita, rapporteur
Prof. M. Payne, rapporteur
Prof. W.-D. Schneider, rapporteur

Lausanne, EPFL
2004

Abstract

This thesis is concerned with the theoretical study of the adsorption of molecules and thin films to single-crystal metal surfaces. First-principles electronic calculations are performed in the framework of density functional theory in the local density approximation (LDA-DFT), by using a plane-wave formalism, norm-conserving pseudopotentials and the supercell method.

For the calculation of the electronic ground state and the subsequent structural optimizations we use the Car-Parrinello technique, modified for the optimal treatment of metallic systems. This method has been improved in this work with the introduction of a Lagrangian formalism and a mathematically conserved constant of motion in the presence of variable fractional occupancies.

Complex mechanisms involving substantial reconstructions of the substrate are elucidated by the study of the C_{60}/Al(111) system. We find that the interaction of C_{60} molecules with the Al(111) surface is predominantly covalent, and that the adsorbates bind optimally to the surface if an Al vacancy is created directly underneath. The removed Al atoms form a (6×6) array of ad-dimers in the interstices of the C_{60} overlayer, to which they strongly bind. Large-scale structural relaxations performed directly on a surface unit cell containing three C_{60} molecules lead to a reconstructed structure with two vacancies and an Al ad-dimer which is significantly more stable than the unreconstructed one. This spontaneous local process, rather than the compression state of the C_{60} overlayer, explains why one C_{60} molecule out of three protrudes from the surface upon reconstruction.

The delicate interplay of many competing factors occurring at the interface between different media is investigated through the study of ultrathin MgO films deposited on the Ag(100) surface. We consider pseudomorphic MgO films of thickness ranging from one to three atomic layers, and we study the evolution of the surface electronic properties as a function of coverage. We find that already three MgO monolayers are sufficient to fully develop the electronic structure of the single-crystal MgO surface, in agreement with experimental findings. We discuss the reliability of the LDA-DFT results for the band structure by using simple 1D models for the dielectric properties of the system.

Version abrégée

Cette thèse concerne l'étude théorique des phénomènes d'adsorption de molécules sur les surfaces cristallines des métaux. Les calculs s'inscrivent dans le cadre de la Théorie de la Fonctionnelle de la Densité dans l'Approximation de Densité Locale (DFT-LDA), en utilisant les pseudo-potentiels, une base d'ondes planes et la technique du super-réseau.

Pour le calcul de l'état fondamental électronique et les relaxations de la structure atomique, nous utilisons la méthode de Car et Parrinello, modifiée pour le traitement optimal des systèmes métalliques. Cette méthode a été améliorée au cours de ce travail par l'introduction d'un formalisme Lagrangien et d'une constante du mouvement mathématiquement conservée en présence d'occupations variables.

Des mécanismes complexes comportant une reconstruction importante du substrat sont approfondis par l'étude du système C_{60}/Al(111). Nous trouvons que l'interaction des molécules de C_{60} avec la surface (111) de l'aluminium est principalement covalente, et que la liaison est optimale lorsqu'une lacune est créée dans la surface d'Al au site d'adsorption. Les atomes d'Al ainsi enlevés vont former un réseau périodique (6×6) dans les interstices de la couche de C_{60}, à laquelle ils se lient fortement. Des relaxations *ab-initio* à grande échelle, effectuées directement sur une cellule unitaire de surface contenant trois molécules de C_{60}, fournissent une structure reconstruite avec deux lacunes et un dimère de Al, qui apparaît beaucoup plus stable par rapport à la structure non reconstruite.

L'équilibre subtil de plusieurs facteurs concurrents qui influence les propriétés de l'interface entre deux matériaux différents à l'échelle atomique est investigué à travers l'étude des films ultra-minces de MgO adsorbés à la surface (100) de Ag. Nous considérons des films de MgO d'une épaisseur allant de une à trois couches atomiques, et nous étudions l'evolution des propriétés électroniques de surface en fonction de la couverture. Nous trouvons que trois monocouches de MgO sont suffisantes pour obtenir une structure électronique pratiquement identique à celle de la surface (100) d'un cristal de MgO pur, en accord avec les résultats des expériences. Nous discutons la fiabilité des résultats LDA-DFT pour la structure de bande en utilisant des modèles 1D qui tiennent compte des propriétés diélectriques du système.

Contents

Chapter 1

Introduction

Small clusters and thin films display peculiar properties which are distinctively different from those of both bulk materials and isolated atoms, and are of current interest for a number of potential applications. Nanoscience is an emerging research field which aims at understanding these properties at a fundamental level and exploiting them in practical applications.

Among the broad variety of systems that are currently investigated, much effort has been devoted to explore the properties of carbon-based nanostructures, novel materials whose building blocks range from small fullerene cages to elongated quasi one-dimensional nanotubes. The latter have many potential applications, e.g. as fibers in strong wires or composites, nano-pipettes to inject proteins into living cells, or electron guns lighting up the phosphor layer on flat-panel displays [6]. The former, and in particular the C_{60} molecule, revealed high-T_C superconductivity upon alkali metal doping [45], and have been successfully used to build the first transistor based on a single molecule as active element [44]. These systems, and their future large-scale industrial applications, rely critically on (i) the macroscopic availability of low-defect and well defined fullerene/nanotube molecules and (ii) their reliable interfacing with larger devices and supporting materials. As a consequence, intensive research is going on to study the structural and electronic modifications of C_{60} when it is deposited on various types of substrates, in particular metallic ones where charge transfer could mimick alkali-metal doping and thus determine superconducting behaviour.

Exciting progress in this research field has been made possible by the extreme accuracy of modern experimental techniques, which are able to control details of the sample at the atomic scale. Scanning tunnelling microscopy (STM) has allowed for a major step forward in the structural and electronic characterization of such systems [10], but is also used for tip-assisted atom-by-atom crafting of nanostructures, and for true molecular engineering by cutting single bonds, deplacing radicals on a surface and inducing chemical reactions between single adsorbed molecules [8]. These techniques are, however, difficult to transfer to large-scale production, being time-consuming and not easily reproducible in the context of industrial processes. Alternative routes are being explored, and the most promising direction relies on the *self-assembly* of adsorbed building blocks into supramolecular structures. Small organic molecules, e.g. nitronaphtalene [3] or PVBA [11], interact rather weakly with a low-index noble metal surface, and are thus highly mobile and free to arrange themselves according to their reciprocal electrostatic interaction. An appropriate control of the self-organizing process may lead to molecular devices such as conducting nanowires or logical gates, which are the basis for a realistic nanocircuit.

It is maybe early to state if and when these techniques will give a concrete alternative to the present silicon-based technology (see e.g. Refs. [9, 7] for more details), but even in current microelectronics there are many issues where nanoscience begins to play a fundamental role. With the ongoing miniaturization, not only the planar extension of a device but also the thickness of insulating layers is becoming smaller and smaller, already reaching in some cases (e.g. magnetic tunnel junctions) the 1-2 nanometer length scale. Up to now, mostly polycrystalline or amorphous oxide films (SiO_x, AlO_x) have been used. However, as size is further reduced, there might be a demand for more homogeneous epitaxial systems with well defined insulating properties. It is well known that below a critical thickness of the order of a few atomic layers the electronic properties of an ultrathin insulating film can be in striking contrast with those of the corresponding bulk crystal. This fact could pose strong physical limits to the current miniaturization trend. On the other hand, an accurate fundamental knowledge of these systems could allow one to exploit the unique behaviour of matter in reduced dimensionality, and thus open new perspectives in the development of oxide heterostructure-based nanodevices.

At the fundamental level, it is important (although not always straigthforward) to make a clear distinction between the properties of small clusters and thin films that are due to their reduced size, and those arising from the interaction with an ideal or defective substrate. Adsorbate-substrate interaction is commonly rationalized into two broad pictures, *physisorption* and *chemisorption*. A molecule is physisorbed when the interaction is weak, of predominant van der Waals nature. This means that the adsorbed species have a rather high mobility and return to gas phase upon annealing at a mild to intermediate temperature. In other cases true chemical bonds are formed, and as a consequence the deposited overlayer is very sensitive to the atomistic details of the underlying substrate. This latter binding is indicated as *chemisorption*, and is often associated to strong perturbation of the electronic structure of the adsorbates, which could even dissociate into constituent atoms and react easily with other species (e.g., catalytic oxidation of CO). Physisorption is appropriate, e.g., for the study of the above mentioned self-assembly processes, and (with particular relevance to this work) for an accurate investigation of deposited insulating thin films. In fact, the substrate acts as an inert support which allows a reliable characterization of the film under study, and the evolution of its intrinsic electronic and structural properties as a function of thickness. The interest in chemisorption, apart from catalysis, stems from the fact that the combined substrate-adsorbate system often displays novel properties, entirely different from those of the constituents. This is e.g. the case of C_{60}, which in the past was thought to be essentially inert, but over the years many studies revealed a rich and unexpected chemistry for its interaction with a variety of crystalline surfaces.

In summary, surface science and nanoscience are two strongly compenetrating research fields. In this study we will consider two systems which are relevant for both. They are complementary to each other and well representative of many of the topics discussed above. We will study the structural and electronic properties of C_{60} deposited on the (111) surface of aluminum, which is a somewhat special case of chemisorption, and ultrathin MgO films physisorbed on the (100) surface of silver, with particular attention to the surface electronic structure and its evolution as a function of film thickness. These studies are entirely based on theoretical *ab-initio* simulations, and the results will be used for the interpretation of available experimental data. For this reason, in the following section we will give a brief overview of the experimental techniques which are relevant in the characterization of these systems. We will try to point out what can be obtained directly from the experiments, and how theoretical results can corroborate these findings, thus contributing to a reliable rationalization of measured quantities.

1.1 Experimental techniques

The scanning tunnelling microscope (STM) provides an unrivaled method for obtaining real space topographs of a surface region, and has become the natural tool for nanoscale investigations. The STM tip is a very versatile instrument, not only for imaging but also for manipulation and spectroscopy at the atomic level. Most importantly, it is sensitive to local variations in the surface electronic structure, and the measured spectra can be readily compared to theoretically predicted local density of states (LDOS) [110]. In many cases molecular or even atomic resolution can be obtained, thus providing invaluable information about atomic equilibrium structures, that can then be used as input for theoretical simulations. Geometric and electronic effects cannot be always disentangled in a straightforward way and some care must be taken when interpreting spectra, especially those collected at high bias voltage, but these are minor issues with respect to the enormous contributions of STM to surface science since its conception. Nevertheless, STM is not the only probing tool and other techniques are routinely used in surface science studies.

Direct information about elemental composition of the sample surface can be obtained by means of Auger photoelectron spectroscopy (APS) and x-ray photoelectron spectroscopy (XPS). These techniques are quite sensitive to the chemical environment of the atoms, and can give direct insight about covalent bonds and substrate-induced rehybridizations in the adsorbed molecules. Shifts in the peak derived from the Al $2p$ core orbitals revealed e.g. the presence of Al adatoms on the surface of Al(111) in the presence of an adsorbed C_{60} overlayer. In the same system, a shift of the C $1s$ derived peak suggested a partial sp^3

rehybridization of the C atoms interacting with the Al substrate [46].

Low-energy electron diffraction (LEED) provides information about the periodic structure of crystal surfaces and the arrangement of adsorbates. It should be considered the surface science equivalent of x-ray crystallography, which is a standard tool in the study of bulk materials. Unlike STM it cannot access *local* structures, but gives an overall picture of the adsorption process, since sharp peaks indicate high-quality commensurate growth, while diffuse or multiple peaks reveal lattice mismatch or some degree of orientational disorder.

Again in the context of C_{60}, very accurate information about the predominant *orientation* of single adsorbed molecules can be obtained by means of x-ray photoelectron diffraction (XPD) techniques [54]. It was applied to C_{60} adsorbed on many substrates, and results were generally in good agreement with molecularly resolved orbital STM images when available [51]. Other techniques are sometimes used to probe surface structure. An accurate analysis of intensity angular distribution (IAD) of Auger photoelectron emission has been applied e.g. to measure surface “rumpling” and equilibrium interlayer spacings in thin MgO films [72], while reflection high-energy electron diffraction (RHEED) has been used to monitor the epitaxial growth of the film [70].

Concerning the electronic structure of surfaces, direct information about quasiparticle spectra and lifetime is readily provided by ultraviolet photoelectron spectroscopy (UPS) and inverse photoemission spectroscopy (IPS). These techniques, especially IPS, are affected by severe sample charging problems when applied to wide bandgap insulators. The study of thin oxide films supported by a metal substrate can solve these problems, but no IPS investigation has been reported yet to our knowledge. Recent UPS studies on thin insulating films tried to answer some long standing issues, e.g. the ionization potential of (100) surfaces for MgO [81], NaCl [13] and LiF [14] was estimated with improved accuracy, but the proximity of the metal can introduce some uncertainity which was tentatively assigned to final state effects [15] and should be carefully investigated to improve reliability. Excited surface states are routinely investigated by electron energy loss spectroscopy (EELS). Excitonic effects should also be taken into account in the interpretation of these latter spectra.

1.2 Theoretical methods and tools

Theoretical simulations are a powerful tool, in some respects complementary to experimental probes, for the investigation of nanoscopic systems. The ingredients of a successful simulation are the basic properties of the constituents, which are integrated numerically to predict the behaviour of the ensemble. Each length scale has its proper description in terms of constituents and laws governing their behaviour, e.g. planets are governed by gravitation, while a bridge is built from elastic materials with well characterized macroscopic properties. At the nanometer scale, atoms are governed by quantum mechanics and a realistic simulation has to treat explicitly electrons and nuclei, except for some particular cases e.g. in materials science where properly adjusted classical interatomic potentials may be sufficient.

In the theoretical study of materials at the atomic scale, density functional theory (DFT) [16] has become more and more widely used. This success stems from a favorable compromise between accuracy, which is comparable in some cases to quantum chemical techniques, and computational efficiency. In the vast majority of realistic surface science problems, which can require relatively large simulation cells, there is practically no alternative to DFT, since more accurate methods are unaffordable with present computer resources. DFT results are quite accurate, and often give quantitative agreement with experimental data. Nevertheless, these results should always be critically discussed, and their validity carefully assessed in each case.

The practical ease, and sometimes even the feasibility, of a simulation depend not only on the underlying theoretical framework, but also on the algorythmic implementation of the formalism. The Car-Parrinello [33] method is a well estabilished scheme that allows for efficient calculation of the electronic ground state, optimization of atomic structures and molecular dynamics simulations. This method experiences well known difficulties when dealing with metals [34], which are the substrates of our systems, and special care must be used to ensure optimal convergence.

A computer code that is perfect *as is* for all applications does not exist to our knowledge, and to some degree one has to tailor it to the specific needs of a given study. This is especially true when dealing with large systems that are at the limit of the currently affordable computational load. The optimization of every single routine is then crucial, while user-friendliness could represent an unneeded complication in the program structure. On the other hand, in some cases one has to do many accurate calculations for relatively small systems. In this case the opposite is true, i.e. the preparation of the input files by the user represents the main “bottleneck”, and an all-purpose easy-to-use package could be extremely useful.

We have therefore dedicated part of our activity to the improvement and optimization of the LAUTREC package, a high-performance Car-Parrinello code for condensed-matter applications. This programming effort was of paramount importance for dealing with the large C_{60}/Al(111) supercells. On the other hand, for the study of the MgO/Ag(100) system, we chose the ABINIT [113] program, which provides a nice and user-friendly interface and thus facilitates substantially the repetitive process of editing the input files.

An accurate theory and optimized algorythms are not sufficient to solve the problems, since the ultimate computing performance is dictated by the hardware quality. The LAUTREC package was used as a test case to assess the performance of several parallel architectures, some of which have been directly used in this work [4].

1.3 Motivation and plan of the thesis

C_{60} adsorption on metals has been intensively studied experimentally [46], but no *ab-initio* investigation has ever been made to our knowledge. There is a rich variety of experimental data (relative to the Al(111) substrate, but also many other metals) which point to exciting and not well understood phenomena at the atomistic level, so we consider this lack of theoretical studies rather surprising. Preliminary calculations on the adsorption geometries and equilibrium energetics of isolated C_{60} on the aluminum surface gave a first insight about the experimental covalent bonding picture. Moreover, they revealed unusual adsorbate-induced substrate relaxations, which were rather promising in sight of a possible explanation for the observed (6×6) reconstruction, which presented many puzzling issues concerning the underlying atomic mechanisms, and motivated us to pursue this route. During early work at IRRMA the LAUTREC program was improved, in particular we removed the limitations to Γ point sampling and to orthogonal supercells, making possible direct investigations of the hexagonal overlayer, and the calculation of the (6×6) supercell with three C_{60} molecules. Thanks to the installation at EPFL of powerful computer servers, there was thus the timely opportunity to perform a thorough study of the C_{60}/Al(111) system and of its reconstruction.

Considering the metallic nature of the substrate, and the recent availability of a practical generalization [38] of the Car-Parrinello method for treating metallic systems, we decided to look for possible optimizations and improvements of the algorythm, in order to reduce the difficulty of the application to such a complex system as C_{60}/Al(111). We proposed (Chapter 3) a fictitious Lagrangian which contains electronic and occupational degrees of freedom, and we demonstrated the performance of this new scheme by performing numerical tests on selected systems. The Lagrangian formulation was not strictly necessary in this work, since we have been mostly concerned with structure optimizations, which are optimally performed within the existing schemes. However, this algorythmic activity has been useful to become acquainted with the method, and for the optimal choice of the simulation parameters.

In Chapter 4 we present the results obtained for the C_{60}/Al(111) system. We discuss the structural, energetic and electronic properties of the full unreconstructed C_{60} overlayer, which is known experimentally to grow commensurate to the substrate in a $2\sqrt{3} \times 2\sqrt{3}R30$ phase. We show that, as a consequence of an unusual steric effect associated with the adsorption bonding geometry, the presence of the C_{60} overlayer drastically reduces the energy cost of adatom-vacancy pair creation in the metal substrate. This can yield a reconstruction mechanism where Al adatoms are produced at the adsorption sites and move to the interstices below the C_{60} overlayer, to which they strongly bind. This spontaneous local process, rather than the elastic strain in the unreconstructed C_{60} overlayer, can explain why one molecule out of three protrudes from the surface upon reconstruction, as we demonstrate by large-scale first-principles structure relaxations performed directly on the (6×6) system.

The work on MgO/Ag(100) was motivated by very interesting experimental data [116] on the evolution of the electronic structure of MgO films as a function of film thickness. The results we obtained in a first investigation were in good agreement with experimental findings [2]. This success, together with some residual issues which were not fully investigated at that stage, motivated us to pursue our study in order to achieve a more complete knowledge of the system. Since then, several experimental and theoretical investigations were reported, which contributed to the improvement of the overall physical picture. However, none of them could give a definitive answer to justify the striking experimental result that even one MgO supported monolayer has an electronic structure which is similar to that of the bulk single-crystal surface. We propose a simple and physically appealing interpretation of this fact, we discuss the possible reasons why LDA is not capable of reproducing quantitative details, and we correlate, as far as possible, the theoretical results with experimental data (Chapter 5).

Before presenting our original results (Chapters 3, 4 and 5), we give in the next Chapter a short introduction to the theoretical foundations of this work, in particular density functional theory (DFT), the Brillouin zone sampling methods, pseudopotentials and basis set representations, and the supercell technique.

Chapter 2

Electronic structure methods

Electrons and nuclei are the "elementary particles" of solid state physics, and their behaviour is described by the laws of quantum mechanics. The goal of the *ab-initio* approaches is to extract information about real physical systems starting from fundamental theory alone, i.e. without using any empirical data as input. This task is not easy to achieve, as the exact Schrödinger equation is far too complex to allow for a direct solution. Despite this inherent difficulty, the *ab-initio* methods are nowadays an invaluable tool in rationalizing experimental data, and in predicting properties of a wide range of materials. The reasons for this success are the existence of reasonable approximations that make the theory tractable by numerical means, the development of efficient algorythms that help optimize the computational effort, and the availability of increasingly powerful workstations and parallel computers thanks to the breathtaking technological progresses. In this chapter we will give a brief introduction to the theoretical tools used in this thesis, while a quick technical overview of their practical implementation is given in the Appendices.

2.1 General assumptions

Most *ab-initio* computational studies rely on two basic assumptions:

- the Born-Oppenheimer (BO) approximation, i.e. the electrons follow adiabatically the motion of nuclei while remaining in their ground state;
- the nuclei are described as classical particles.

This is justified by the relatively large mass of the nuclei compared to the electrons [1]. We can now treat the electronic structure and the ion dynamics

---

[1] There are many physical phenomena where this *Ansatz* does not hold very well, i.e. systems involving light atoms (e.g. hydrogen bonds), or nonadiabatic transitions in chemical

separately; in the reminder of this Chapter we will focus on the electrons only. Neglecting relativistic effects, the ground-state wavefunction of the N electrons in the Coulomb potential of the nuclei (which are represented as point charges) is given by the lowest energy solution of the Schrödinger equation

$$\hat{H}\Psi(\mathbf{r}_1, ..., \mathbf{r}_N, \sigma_1, ..., \sigma_N) = E\Psi(\mathbf{r}_1, ..., \mathbf{r}_N, \sigma_1, ..., \sigma_N), \tag{2.1}$$

where $\mathbf{r}_i$ and σ_i indicate space and spin coordinates of the electrons, and the hamiltonian $\hat{H}$ depends parametrically on the ionic positions $\mathbf{R}_\nu$:

$$\hat{H} = -\frac{1}{2}\sum_i \nabla_i^2 - \sum_i V_{ion}(\{\mathbf{R}_\nu\}; \mathbf{r}_i) + \frac{1}{2}\sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{2.2}$$

The N-particle wavefunction Ψ depends upon $3N$ real variables, and due to the electron-electron repulsion term it cannot be factorized as a product of one-particle orbitals. Therefore, apart from small molecules and light atoms (with $N < 10$) the numerical solution of Eq. 2.1 is very difficult if not impossible. We will see in the next section how Density Functional Theory provides an alternative and elegant approach to this problem.

2.2 Density Functional Theory

In their seminal paper Hohenberg and Kohn [16] showed rigorously that the total energy of a many-electron system in an external potential $V_{ion}(\mathbf{r})$ (we will drop in the following the parametric dependence of the potential on the atomic coordinates) is uniquely determined by the ground-state electronic density $n(\mathbf{r})$:

$$E[n(\mathbf{r})] = \int V_{ion}(\mathbf{r})n(\mathbf{r})d\mathbf{r} + F[n(\mathbf{r})], \tag{2.3}$$

where F is a universal functional, i.e. it does not depend on the external potential. Furthermore, a variational principle shows that $E[n(\mathbf{r})]$ has an absolute minimum when $n(\mathbf{r})$ is the density corresponding to the exact ground state. This allows for the replacement of the many-particle wavefunction with the ground state density as the basic variable of the problem, with a considerable gain in efficiency. Unfortunately, the functional F is unknown, and contains implicitly all the complicated quantum many-body effects. A major step toward a practical formulation has been taken by Kohn and Sham [17] one year later. They proposed a decomposition of the charge density into a set of one-particle orthonormal orbitals:

$$n(\mathbf{r}) = \sum_{i=1}^{N} |\psi(\mathbf{r})|^2, \tag{2.4}$$

and wrote the functional $F[n]$ in the form

$$F[n] = T_0[n] + \frac{1}{2}\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} dr dr' + E_{xc}[n] \tag{2.5}$$

---

reactions, but their discussion falls outside the scope of this work.

where $T_0[n]$ is the kinetic energy of the non-interacting orbital set (this set is uniquely chosen as the minimum of T_0 given the condition 2.4) and the second term is the classical (Hartree) electrostatic term. E_{xc} is the so-called *exchange and correlation* energy, and is implicitly defined by the above expression. The rationale behind this apparently arbitrary decomposition is to minimize the unknown part of the problem by subtracting away some reasonable contributions that are easy to calculate. $E_{xc}[n]$ turns out to be much less important in magnitude than the original $F[n]$, and therefore a better starting point for introducing the approximations, as we will see in the next section.

We can now consider the total energy as a functional of the orbital set $E[\psi_i(\mathbf{r})] = E[n]$, and introduce appropriate Lagrange multipliers for the orthonormality constraint:

$$G = E[\psi_i] + \sum_{ij} \Lambda_{ji} \langle \psi_i | \psi_j \rangle. \tag{2.6}$$

At the minimum G is stationary with respect to first order variation of the orbitals. By imposing this condition, and by choosing for ψ_i the representation where the Λ matrix is diagonal, we obtain a set of single-particle Schrödinger-like equations:

$$\big[-\frac{1}{2}\nabla^2 + V_{ion}(\mathbf{r}) + V_H(\mathbf{r}) + V_{XC}(\mathbf{r}) \big] \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}), \tag{2.7}$$

where the Hartree (H) and exchange-correlation (XC) potentials are defined respectively as:

$$V_H(\mathbf{r}) = \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} dr', \qquad V_{XC}(\mathbf{r}) = \frac{\delta E_{XC}[n]}{\delta n(\mathbf{r})},$$

and ϵ_i are the eigenvalues of Λ. These are the Kohn-Sham (KS) equations. The potential depends on the orbitals through the charge density, and therefore these equations must be solved iteratively until self-consistency is achieved.

In principle, the eigenvalues ϵ_i don't have a direct physical meaning, since they enter the formalism as Lagrange multipliers that ensure the orthogonality of the orbitals. However, it has become standard practice to interpret the ϵ_i as estimates of excitation and electron removal energies, and to compare them with experimental spectra and band structures. Green function approaches are indeed better suited to the study of one-particle excitations in the correlated electron gas [23], but these techniques can be computationally cumbersome, and apart from some notable exceptions [24] they have been applied mainly to bulk systems, so that in the vast majority of cases there is no simple alternative to Kohn-Sham eigenvalues.

One should notice that the KS eigenvalue corresponding to the highest occupied state has real physical meaning [25], at variance to the whole rest of the spectrum, and corresponds to the exact ionization potential (IP). Furthermore, in simple metals the band structure in the neighborhood of the Fermi level is reproduced quite accurately [26], and this gives a good justification for discussing electronic properties on the basis of KS formalism. In atoms with discrete energy levels and extended systems with an energy gap this agreement breaks down [26], in particular for the unoccupied states which appear too low in energy with respect to true quasiparticles. Still, even in insulators, the internal structure of separate KS bands is relatively accurate, and the discrepancy often consists in a rigid upward shift of the unoccupied states relative to the valence bands, which is known in the literature as "scissor operator" [27].

The shortcomings which are inherent to the KS approach are not easy to distinguish from the errors introduced by commonly used approximations for E_{XC} (see following sections). A detailed discussion of this debated issue would fall outside the scope of our work, and here we will limit ourselves to notice that in insulators both features introduce appreciable errors, which should be carefully considered, especially when dealing with wide band gap materials.

2.3 Local Density Approximation

The DFT described above is formally exact for the ground state, but it is useless for practical purposes, because an explicit expression for the exchange and correlation functional does not exist. In order to make a concrete use of DFT, an approximation of E_{XC} is therefore needed. The most common and simplest approach is to assume that, given an inhomogeneous charge distribution $n(\mathbf{r})$, the exchange-correlation energy density per electron $\epsilon_{XC}(\mathbf{r})$ is locally equivalent to the same quantity evaluated in an homogeneous electron gas ϵ_{XC}^{hom} of density $n(\mathbf{r})$:

$$E_{LDA}^{XC}[n(\mathbf{r})] = \int n(\mathbf{r}) \epsilon_{XC}^{hom}(n(\mathbf{r})) dr. \tag{2.8}$$

The XC energy of the homogeneous electron gas is known analytically in the high density limit, and has been calculated by means of Quantum Monte Carlo techniques at low and intermediate densities. These results have been interpolated and parametrized by several authors. In this work we have chosen the parametrization by Perdew and Zunger [21], which is based on the QMC results of Ceperley and Alder [20].

This approximation, quite questionable at a first glance, is surprisingly accurate, even in systems characterized by strong inhomogenities in the electronic charge density. The reasons for this success are not yet fully understood, and intensive research is still going on on this subject.

Nevertheless, there are some cases where LDA does not give satisfactory results, and many attempts have been made to improve it. The Generalized Gradient Approximation (GGA) has become quite popular, because in many cases it gives binding energies that are closer to experiment than LDA, without any major computational overhead. Still, there are cases where LDA gives an overall better description, so a comparison between the two functionals is often needed to understand which one is best suited for a particular system. It is important to notice that both functionals give a local or semilocal description of the exchange and correlation in a many-electron system. This implies that they cannot describe properly truly non-local quantum effects, such as the Van der Waals bonding, or the classical image potential in the proximity of a metallic surface. Systems where the physics is dominated by these interactions should be treated with special care, as we will discuss extensively in Chapters 4 and 5.

2.4 Crystalline periodicity and Brillouin-zone integrations

Many properties of crystalline systems depend only on the bulk structure and composition, and not on the terminating surfaces. To study these properties it is customary to consider periodic infinite crystals, where the Bloch theorem and standard band structure theory can be applied. Each eigenstate of the infinite solid can be then written as a product of an exponential times a function that has the same periodicity as the crystal:

$$\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{kr}} u_{n\mathbf{k}}(\mathbf{r}), \quad u_{n\mathbf{k}}(\mathbf{r}+\mathbf{R_i}) = u_{n\mathbf{k}}(\mathbf{r}) \tag{2.9}$$

where $\mathbf{R_i}$ is a lattice vector, and $\mathbf{k}$ is a reciprocal space vector in the Brillouin zone (BZ). Since the $u_{n\mathbf{k}}$ are periodic, the calculation can be restricted to a single unit cell of the whole solid. Moreover, they can be represented on a discrete plane wave basis set corresponding to the crystal periodicity:

$$u_{n\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{n\mathbf{k}}(\mathbf{G}) e^{i\mathbf{Gr}} \tag{2.10}$$

In practice this sum is restricted to a finite number of terms by retaining only the plane waves with kinetic energy below a certain cutoff E_{cut}. A cutoff as low as possible is very desirable, in order to speed up the actual calculation, but this is not always possible without compromising the precision of the results. The optimal choice of this parameter depends mainly on the atomic species under study, as we will discuss in the following section.

Every physical observable should be evaluated as an integral over the BZ:

$$\langle \hat{A} \rangle = \int_{BZ} d\mathbf{k} \sum_n f_{n\mathbf{k}} \langle \psi_{n\mathbf{k}} | \hat{A} | \psi_{n\mathbf{k}} \rangle, \tag{2.11}$$

where $f_{n\mathbf{k}}$ is the occupancy of the Bloch orbital. This integral can be approximated by a weighted discrete sum:

$$\int_{BZ} d\mathbf{k} \rightarrow \sum_{\mathbf{k}} w_{\mathbf{k}} \tag{2.12}$$

Special point theory and the use of point symmetries allow one to greatly reduce the number of independent $\mathbf{k}$-points actually needed in the calculation.

There is an additional difficulty associated with metals, which possess partially filled bands near the Fermi level. The shape of the Fermi surface in reciprocal space is often quite complicated, and the abrupt step function that separates occupied from unoccupied states leads to poor convergence with the number of **k**-points used. To circumvent this problem, an artificial finite-temperature thermodynamic distribution in the orbital occupations is often introduced. The zero temperature total energy functional is then replaced by the free energy, by adding a suitable entropy term depending on the orbital occupation numbers, which become independent variables of the problem and can assume fractional values. The most natural choice for the distribution function is the Fermi-Dirac statistics, which is dictated by basic principles of quantum statistical mechanics. However, as the finite temperature is not a physical feature but merely a computational tool, one is free to choose other analytic forms for the entropy [2].

2.5 Pseudopotentials

In the well known picture of elementary chemistry, the electrons in free atoms are organized in energy "shells", the inner orbitals (core) being tightly bound to the nucleus and hardly influenced by the interaction with neighbouring species. On the other hand, the outer (valence) electrons lay at higher energy, and are almost entirely responsible for the physical properties of the solid. This insight has lead to the development of pseudopotentials, which replace the nucleus and the "frozen" core with their combined effective influence on the valence orbitals. The crystalline Hamiltonian then needs only to describe the electrons that are most relevant in the study, with a substantial computational saving for complex systems. A modern pseudopotential is constructed *ab-initio*, i.e. the generation normally starts with an all-electron calculation for the ground-state of the isolated atom. Then, the pseudopotential is constructed according to the following criteria:

- the energy eigenvalues for the pseudoatom match exactly the all-electron values;
- the pseudowavefunctions match exactly the all-electron orbitals beyond a certain cut-off radius r_{cut} and have no radial nodes;
- the scattering properties of the pseudopotential are indistinguishable from those of the all-electron system over a wide range of energies;
- the pseudopotential has good convergence properties with respect to the size of the basis set used in the crystalline environment.

The first two requirements cannot in general be fulfilled unless a *non-local* pseudopotential is used, i.e. the radial form is different for each angular momentum component. The third requirement, which is a necessary condition for *transferability*, sets an upper limit on the choice of r_{cut}, which in principle should be taken small enough that the core spheres of neighbouring atoms do

---

[2] The finite electronic temperature can have a real physical meaning, and can be relevant e.g. when simulating the canonical ensemble for a metallic system. The formal extension of DFT to finite temperatures has been formulated by Mermin [19].

not overlap. The fourth requirement concerns (i) the choice of the mathematical form for the non-local operator and (ii) the optimization of the remaining degrees of freedom in order to ensure economy with respect to the basis set used in the crystalline calculations.

Concerning (i), we notice that a literal implementation of the angular non-locality would imply the calculation of projection operators of the form

$$\hat{p}_{lm} = |lm\rangle\langle lm| = \sum_n |nlm\rangle\langle nlm| \tag{2.13}$$

which is rather expensive. The commonly used strategy is to approximate the above infinite sum by retaining only *one* radial orbital, and to write the non-local operator in a fully separable form:

$$\hat{V}_{lm} = \frac{|V_{lm}(\mathbf{r})\phi(\mathbf{r})\rangle\langle V_{lm}(\mathbf{r})\phi(\mathbf{r})|}{\langle\phi(\mathbf{r})|V_{lm}(\mathbf{r})|\phi(\mathbf{r})\rangle} \tag{2.14}$$

The summation on the quantum number l is truncated in practice to a maximum value $l_{max} = 1, ..., 3$ depending on the atomic specie under consideration, and a further gain in efficiency is obtained by using a *local* potential for one angular momentum channel, so that $V_{lm}(\mathbf{r})$ is redefined in the above equation as $V_{lm}(\mathbf{r}) - V_{LOC}(\mathbf{r})$. This is the Kleinman-Bylander [18] representation, which is nowadays the standard implementation of norm-conserving pseudopotentials.

For the specific needs of the plane-wave basis set, the residual freedom is exploited to ensure optimal *smoothness*, i.e. the pseudoorbitals are well represented by a limited energy cut-off in the actual calculation. Nevertheless, for many atomic species it is not possible to obtain smooth pseudopotentials, e.g. second row elements have inherently sharp $2s$ and $2p$ orbitals, and transition metals have localized d states. In these cases a high plane wave cutoff is unavoidable, unless resorting to alternative techniques such as the Vanderbilt [29] ultrasoft scheme or the Blöchl [30] PAW formalism.

In this work we will use norm-conserving pseudopotentials calculated within the Troullier and Martins [22] scheme. These pseudopotentials have been generated with the fhi98PP [56] code, and carefully tested for transferability. The use of non-linear core corrections [28] was necessary to correctly reproduce the structural properties of highly ionic materials (Chapter 5).

2.6 Study of surface systems

A formalism based on Periodic Boundary Conditions seems at first sight to be useless in the study of surfaces and interfaces, which are nonperiodic in the direction perpendicular to the interface plane. This problem can be solved by means of the *supercell* technique, i.e. by considering a thin slab instead of a semiinfinite solid. By repeating identical images of the slab in the z direction, and separating them by a vacuum layer, we can restore the 3D periodicity of the system, and represent it in the same formalism already developed for crystalline bulk systems. The thickness of the slab of course must be large enough to ensure that its surface properties are representative of the semiinfinite system. Moreover, the vacuum layer must be thick enough to avoid unwanted interactions between the repeated images. The convergence of the numerical results with respect to these two additional parameters must be checked carefully in each case.

Chapter 3

Ab-initio molecular dynamics

In the first chapter we briefly presented state-of-the-art methods for calculating the total energy of a many-electron system in the external potential of fixed ions. In this chapter we show how the ionic degrees of freedom are dealt with, focusing on the case of metallic systems.

3.1 Hellmann-Feynman forces

The quantum total energy of the electronic ground state as a function of the ion positions can be thought as the *ab-initio* effective potential acting on a system of classical ions. The forces are the derivative of the energy with respect to the cartesian coordinates of the ions:

$$\mathbf{f}_\nu = \frac{dE}{d\mathbf{R}_\nu} = \frac{\partial E}{\partial \mathbf{R}_\nu} + \sum_i \frac{\partial E}{\partial \psi_i}\frac{\partial \psi_i}{d\mathbf{R}_\nu} + \sum_i \frac{\partial E}{\partial \psi_i^*}\frac{\partial \psi_i^*}{d\mathbf{R}_\nu} \tag{3.1}$$

The Hellmann-Feynman theorem allows one to simplify this expression. Since $\frac{\partial E}{\partial \psi_i^*} = \hat{H}\psi_i$, if ψ_i are eigenstates of $\hat{H}$ it is easy to show that the second and third term vanish. Therefore the force can be written as an expectation value over the ground state orbitals:

$$\mathbf{f}_\nu = \sum_i \langle \psi_i | \frac{dV_{ion}(\mathbf{r})}{d\mathbf{R}_\nu} | \psi_i \rangle \tag{3.2}$$

The most natural technique for performing molecular dynamics simulations is then to integrate directly the Newtonian dynamics generated by these forces. This method, known as "quench and go" technique has the serious drawback that one has to minimize accurately the electronic total energy for every ionic position along the trajectory. This procedure is quite expensive, and the HF forces are very sensitive to small deviations from the exact instantaneous ground state (first order, while the total energy is only second order). Still, the method is widely used because of its simplicity and robustness, and with some care one can avoid in many cases the most serious bottlenecks.

3.2 The Car and Parrinello method

The most impressive breakthrough in the history of *ab-initio* molecular dynamics has been made in 1985 by Car and Parrinello [33]. Instead of quenching the electronic structure at every step, they promoted the wavefunction coefficients to classical degrees of freedom, which undergo a *fictitious* dynamics in the potential defined by the total energy functional. For that purpose, they introduced a mass and a kinetic energy term for these new variables, and wrote a unified Lagrangian for both the ionic and electronic part:

$$\mathcal{L}_{CP} = \sum_{i=1}^{N} \mu \langle \dot{\psi}_i | \dot{\psi}_i \rangle + \sum_{\nu=1}^{n} \frac{1}{2} M_\nu \dot{\mathbf{R}}_\nu^2 - E[\{\psi_i\}; \{\mathbf{R}_\nu\}] \tag{3.3}$$

The simulation starts with fixed ionic positions and a set of trial electron wavefunctions. First of all, the Eulero-Lagrange (EL) equations of motion (e.o.m.) for the electrons are integrated with an appropriate "friction" term until the ground state is reached. Then the ionic and electronic degrees of freedom are left free to evolve at the same time. If certain conditions are satisfiedthe ions describe the correct physical trajectory and the electrons follow them adiabatically while remaining close to the instantaneous ground state. The electronic structure is indeed never at the ground state, but "oscillates" around it so that the first order errors in the forces are averaged out during the motion of the ions. This technique allows a substantial reduction of the computational effort during a simulation, as the exact ground state needs to be calculated only once at the beginning of the procedure.

3.3 Why does it work?

Let's look in deeper detail to the conditions that must be fulfilled for an efficient application of the CP method [1]. The adiabatic separation between ionic and electronic motion is warranted only if the characteristic frequencies of the two subsystems are disjoint and well spaced apart. The ionic frequencies are well known and are determined by the physical properties of the material. The electronic frequencies need to be higher, and the lower boundary of their range is determined by the fake mass and by the lowest restoring force that keeps the electrons close to the ground state. To derive explicitly this restoring force we start with a base of eigenstates of the KS Hamiltonian:

$$\hat{H} |\phi_i\rangle = \epsilon_i |\phi_i\rangle, \tag{3.4}$$

---

[1] We will follow the derivation given in [34]

and we write the instantaneous set of CP orbitals as a small first-order perturbation of the N lowest eigenstates $|\phi_i\rangle$:

$$|\psi_i\rangle = \sum_{k=1}^{\infty}(\delta_{ik} + \alpha_k^{(i)})|\phi_k\rangle, \qquad \alpha_k^{(i)} \ll 1. \tag{3.5}$$

For the sake of generality, we will associate to each orbital its occupation number f_i, and its own fictitious mass μ_i. The e.o.m. for the electrons can then be written:

$$\mu_i\ddot{\psi}_i = -f_i\hat{H}_{KS}|\psi_i\rangle + \sum_j \Lambda_{ji}|\psi_j\rangle \tag{3.6}$$

where Λ_{ji} are the Lagrange multipliers that ensure orthonormality constraints. They can be written explicitly by differentiating twice both sides of the orthogonality equation $\langle\psi_i|\psi_j\rangle = \delta_{ij}$:

$$\langle\psi_i|\ddot{\psi}_j\rangle + \langle\ddot{\psi}_i|\psi_j\rangle + 2\langle\dot{\psi}_i|\dot{\psi}_j\rangle = 0 \tag{3.7}$$

By using the expression (3.6) for $|\ddot{\psi}_i\rangle$ and discarding the squared velocity term which is second order we obtain:

$$\Lambda_{ij} \simeq \frac{\mu_j f_i + \mu_i f_j}{\mu_i + \mu_j} H_{ij}, \tag{3.8}$$

and we can write the generalized force acting on the i-th orbital as follows:

$$F_i = -f_i\hat{H}_{KS}|\psi_i\rangle + \sum_{j=1}^{N}\frac{\mu_j f_i + \mu_i f_j}{\mu_i + \mu_j}H_{ji}|\psi_j\rangle + O(\alpha^2) \tag{3.9}$$

Now we replace the $|\psi_i\rangle$ by the expression (3.5):

$$F_i = -f_i\sum_{k=1}^{\infty}(\delta_{ik} + \alpha_k^{(i)})\epsilon_k|\phi_k\rangle +$$

$$+\sum_{l=1}^{N}\frac{\mu_l f_i + \mu_i f_l}{\mu_i + \mu_l}\sum_{k=1}^{\infty}(\delta_{lk} + \alpha_k^{(l)})\epsilon_k\sum_{k'=1}^{\infty}(\delta_{lk'} + \alpha_{k'}^{(l)})|\phi_{k'}\rangle + O(\alpha^2). \tag{3.10}$$

By grouping together the summations from $N+1$ to ∞ and those from 1 to N, and by discarding all terms that are squared or cubic in the α coefficients we obtain:

$$F_i = -f_i\sum_{k=N+1}^{\infty}\alpha_k^{(i)}(\epsilon_k - \epsilon_i)|\phi_k\rangle +$$

$$+\sum_{l=1}^{N}\Big(-f_i\alpha_l^{(i)}\epsilon_l + f_i\alpha_l^{(i)}\epsilon_i + \frac{\mu_l f_i + \mu_i f_l}{\mu_i + \mu_l}\alpha_i^{(l)}\epsilon_i + \frac{\mu_l f_i + \mu_i f_l}{\mu_i + \mu_l}\alpha_l^{(i)}\epsilon_l\Big)|\phi_l\rangle + O(\alpha^2). \tag{3.11}$$

Due to the orthogonality constraint the relation $\alpha_i^{(l)*} + \alpha_l^{(i)} = 0$ must be satisfied for all $i, l = 1, ..., N$, and the expression can be simplified to its final form:

$$F_i = -f_i \sum_{k=N+1}^{\infty} \alpha_k^{(i)} (\epsilon_k - \epsilon_i) |\xi_k\rangle - \sum_{l=1}^{N} \frac{\mu_i}{\mu_i + \mu_l} (f_i - f_l)(\epsilon_l - \epsilon_i) \alpha_l^{(i)} |\xi_l\rangle + O(\alpha^2). \tag{3.12}$$

Standard CP dynamics is always performed on systems where all orbitals are fully occupied and a unique fake mass is used ($f_i = 2, \mu_i = \mu$). In that case, the second sum in the above expression is zero as we expected, because the Hamiltonian is invariant with respect to unitary rotations among the occupied states, and the associated characteristic frequency is zero. The first sum gives rise to frequencies $\omega_{ik} \sim \sqrt{\frac{\epsilon_k - \epsilon_i}{\mu}}$, where $i \leq N$ and $k \geq N+1$. The minimal value is therefore determined by the gap in the eigenvalue spectrum between occupied and unoccupied states. In semiconducting and insulating systems the adiabatic separation can always be achieved by a proper choice of the fake mass, but in metals this is not possible because of the vanishing electronic gap. We will treat this point extensively in the next section.

We notice here that the performance of a Car Parrinello simulation is strongly influenced by the cutoff energy in the plane wave basis. This parameter determines the *maximal* frequency of the electrons, and therefore poses an upper limit on the timestep that can be adopted in the integration of the e.o.m. [2] A strategy to overcome this problem is the mass preconditioning method proposed by Tassone *et al.* [35].

Finally, for very large systems, the long-range electrostatic interaction can cause strong oscillations in the charge density and thus in the self-consistent potential (*charge sloshing*). This again poses serious limits on the timestep, and a partial solution has been proposed [35] which consists in temporarily switching to an exponentially decaying Yukawa-like interaction.

3.4 Metallic systems

Metals have a vanishing electronic gap, and therefore a straightforward application of the CP method is difficult, because the adiabatic separation of the ionic and electronic subsystems is not possible. It is true that in practical calculations the gap is never exactly zero even when dealing with metals, because of the discreteness of the k-point sampling, but there is an additional problem that makes things even worse. In fact, during ion dynamics, a metallic system is in general characterized by frequent level crossings near the Fermi level. This means that, in the energy band-structure diagram of a given system, ϵ_{n+1} (lowest unoccupied state) crosses ϵ_n (highest occupied state). After the crossing, the n state lies in a "unstable equilibrium" condition. One can imagine a simple pendulum oscillating near its energy minimum, when suddenly the gravitational field

---

[2] We note that the efficiency of an algorythm is the simulation time that can be covered with a fixed number of operations, and thus it is directly proportional to the timestep used.

is reversed. The consequence is catastrophic, since the small-oscillation regime breaks down and the fictitious electronic subsystem accumulates a large amount of kinetic energy.

Many solutions have been proposed in order to circumvent these serious shortcomings. In the following we will present the evolution of the ideas in an historical order.

Blöchl A solution for keeping the electrons reasonably close to the ground state has been proposed by Blöchl and Parrinello [31]. They suggested to couple a Nosé [32] thermostat to the electronic subsystem in order to keep their kinetic energy low. To avoid the cooling down of the ions caused by the energy transfer a second thermostat on the ion dynamics is then needed. This solution, despite its successful application in a number of studies, is somewhat unsatisfactory, because (i) microcanonical dynamics is no more possible, (ii) an accurate tuning of the thermostat is needed for each system and simulation temperature, (iii) in cases where the energy transfer is very large the overall quality of the ionic trajectories is questionable.

Grumbach A more general problem is that the Car Parrinello method does not allow variable or fractional occupation numbers, leading to poor k-point sampling in metallic systems. The first attempt to generalize the formalism for variable occupancies has been done by Grumbach *et al.* [37]. They introduced a new set of dynamical variables with the dimension of energy ξ_i and, given a distribution function $f(\epsilon)$, they defined the occupation numbers as $f_i = f(\xi_i)$. The fictitious Lagrangian then reads:

$$\mathcal{L}'_{CP} = \sum_{i=1}^{N} \mu_i \langle \dot\psi_i | \dot\psi_i \rangle + \frac{1}{2} \sum_i Q_i \dot\xi_i^2 + \sum_{\nu=1}^{N} \frac{\mathbf{P}_\nu{}^2}{2M_\nu} - F[\{\psi_i\};\{\xi_i\};\{\mathbf{R}_\nu\}] \tag{3.13}$$

where Q_i are the masses associated to the new variables. The holonomic constraint of charge conservation $\sum_i f_i = N$ must now be enforced during the evolution of the ξ_i. The total energy E has been replaced by the free energy $F = E - TS$, where the entropy S depends on the occupation numbers only. In principle, the functional form of S determines uniquely the shape of the distribution function $f(\epsilon)$ that minimizes the total free energy, but the f_i are difficult to treat directly as dynamic variables, and that is the reason why the distribution function is parametrized explicitly, and the ξ_i are used instead of the f_i. The e.o.m. for the ψ_i are the same as in (3.6). The e.o.m. for the ξ_i are

$$\frac{\partial F}{\partial \xi_i} = \frac{\partial E}{\partial \xi_i} + T\frac{\partial S}{\partial \xi_i} + \lambda \frac{\partial f(\xi_i)}{\partial \xi_i}, \tag{3.14}$$

where λ is the Lagrange multiplier introduced to enforce total charge conservation. Using the notation $g_i = \frac{\partial f(\xi_i)}{\partial \xi_i}$ and noticing that in general $\frac{\partial S}{\partial f_i} = \xi_i/T$ we can write the e.o.m. as:

$$Q_i \ddot\xi_i = g_i(H_{ii} - \epsilon_i + \lambda) \tag{3.15}$$

We notice here that the f_i and g_i can assume a wide variety of values, ranging from very small (in the upper part of the spectrum) to values of the order of unity near the Fermi level. The generalized forces derived in 3.6 and 3.15 are proportional to these values, and can therefore be vanishing for states that are high in energy. As a consequence, in order to avoid slow frequency modes, the masses μ_i and Q_i must be chosen accurately. The most reasonable choice is:

$$\mu_i = \mu f_i, \qquad Q_i = -Q g_i, \tag{3.16}$$

because in this case the f_i and g_i coefficients cancel out in the leading terms of both sides of equations 3.6 and 3.15. However, this implies that the masses are no more constant, and their variation must be taken into account in the derivation of the Eulero-Lagrange equations. The authors chose to avoid this complication by enforcing 3.16 only from time to time, and keeping the masses constant during normal evolution. The resulting equations are:

$$\mu f_i \ddot{\psi}_i = -f_i \hat{H}_{KS} \psi_i + \sum_j \Lambda_{ji} \psi_j \tag{3.17}$$

$$Q g_i \ddot{\xi}_i = g_i (H_{ii} - \epsilon_i + \lambda) \tag{3.18}$$

The practical implementation of these equations gave poor results. In fact, the energy transfer during molecular dynamics runs turned out to be much more serious than in standard CP dynamics performed on the same system. This failure is easy to understand by considering that the finite temperature (FT) functional is no longer invariant with respect to unitary transformations of the ψ_i. In particular, the free energy is at the minimum if and only if the ψ_i are exact eigenvectors of the KS Hamiltonian, while in the zero temperature (ZT) functional any linear combination was equivalent. Looking again at the first-order perturbative result (Eq. 3.5), we can easily see that if $f_i \neq f_j$ for some i, j the corresponding "rotational" mode has a characteristic frequency $\omega_{ij} \sim \sqrt{(f_i - f_j)(\epsilon_j - \epsilon_i)}$. This frequency is arbitrarily low for states that lie close in energy. Thus many low-frequency modes make the fictitious electronic dynamics ill-conditioned, and seriously limit the applicability of this method.

Marzari A few years later Marzari, Vanderbilt and Payne [36] made a step forward, and treated occupations and rotations as part of the same problem. They replaced the occupation *numbers* with an occupation *matrix* f_{ij}. The charge density is then defined as

$$n(\mathbf{r}) = \sum_{ij} f_{ji} \psi_i^*(\mathbf{r}) \psi_j(\mathbf{r}) \tag{3.19}$$

The functional to be minimized is

$$A[\{\psi_i\}; \{f_{ij}\}; T] = \sum_{ij} f_{ji} \langle \psi_i | \hat{T} + \hat{V}_{ext} | \psi_j \rangle + E_{HXC}[n] - TS[\{f_{ij}\}] \tag{3.20}$$

This functional is invariant with respect to a unitary transformation of *both* orbitals and occupation matrix. This can be verified by letting $\mathbf{f} \rightarrow \mathbf{f}' = \mathbf{U}\mathbf{f}\mathbf{U}^\dagger$ and $|\psi_i\rangle \rightarrow |\psi'_i\rangle = \sum_j |\psi_j\rangle U_{ji}$. This important property of the functional A allows one to define a new projected functional G, which depends on the $|\psi_i\rangle$ only:

$$G(\{\psi_i\}, T) = \min_{\{f_{ij}\}} A[\{\psi_i\}, \{f_{ij}\}; T] \tag{3.21}$$

In the G functional the occupation matrix appears as an implicit, internal variable. This result is very appealing, because it eliminates the need for the complicated dynamical treatment of the occupancies, and at the same time it cures all the conditioning problems caused by subspace rotations. G is indeed invariant with respect to unitary transformations of the orbitals, and therefore it could be successfully used as total energy functional in a standard CP dynamics. However, the implicit treatment of the occupation matrix needs a nested minimization cycle to be performed on the f_{ij} variables for each CP timestep. This fact implies a substantial increase in the computational load. In fact the optimization of the f_{ij} has to be done self-consistently, and typically it requires two to four internal cycles, whose numerical load is comparable to that of a standard CP step. For this reason, the use of G in the framework of the CP method is somewhat unsatisfactory, and this formalism has been applied only to conjugate-gradient minimizations and molecular dynamics applications of the “quench and go” style. Nevertheless, this formalism has the great merit of treating simultaneously occupations and rotations as part of the same problem.

Vandevondele Following the invariant formalism described above, Vandevondele and De Vita [38] proposed a *dynamical* scheme for the evolution of electronic and occupational variables:

$$\begin{aligned} Q\ddot{\xi}_{ij} &= -(\xi_{ij} - h_{ij} - \lambda\delta_{ij}) && (3.22) \\ \mu f_{ll}|\ddot{\psi}_l\rangle &= -f_{ll}\hat{H}|\psi_l\rangle + \sum_k \Lambda_{kl}|\psi_k\rangle && (3.23) \end{aligned}$$

Eq. (3.23) looks identical to the *Grumbach* expression, but it is actually rather different, since the equations are now invariant with respect to subspace rotations, and are expressed in the reference frame where $\mathbf{f}$ is diagonal only for convenience. Eq. (3.22) is the equation of a simple harmonic oscillator with a single characteristic frequency $\omega_Q = \sqrt{1/Q}$ that keeps the matrix elements of ξ close to those of the hamiltonian matrix. In particular, the dynamical motion of ξ keeps the wavefunctions ψ close to the *eigenvectors* of $\hat{H}$. This fact has an important consequence within the perturbative analysis of Eq. (3.5), since it sets a constraint on the possible first order perturbations which the wavefunctions can effectively undergo. In particular, in the ideal case where $[\xi, h] = 0$ (which corresponds to the limit of very low Q), one has $\alpha_k^{(i)} = 0$ for all $i, k = 1, ..., m$ with m corresponding to the total number of orbitals ($m > N/2$, since in a finite-temperature formulation one must include in the calculation several unoccupied states with energy above E_F). This means that the rotational forces (second sum on the right hand side of Eq. (3.5)) are suppressed and the corresponding frequencies are compressed to zero; expression (3.5) then reads:

$$F_i = -f_i \sum_{k=m+1}^{\infty} \alpha_k^{(i)} (\epsilon_k - \epsilon_i)|\xi_k\rangle + O(\alpha^2). \tag{3.24}$$

Therefore, the characteristic frequencies of the electronic fictitious dynamics are, as in the CP zero-temperature method for insulators, $\omega_{ik} \sim \sqrt{\frac{\epsilon_k - \epsilon_i}{\mu}}$, but here $i \leq m$ and $k \geq m+1$. To see the impact of this result on the dynamics, we note that the total energy and the particle density are almost completely determined by the most occupied wavefunctions, i.e. those with $\epsilon_i < E_F$. For all the frequencies of oscillation of these wavefunctions around the minimum we find that $\omega > \sqrt{E_{AG}/\mu}$, where E_{AG} is an "artificial" energy gap corresponding to the range of the eigenvalue spectrum where the occupancies are vanishing. We remark that E_{AG} is not a physical property of the system, but a parameter that can be increased arbitrarily by including more empty states in the simulation. When enough electronic states are included, the problem of energy transfer is removed for all practical purposes, since all slowly oscillating states are negligibly occupied, and therefore do not couple significantly to the ionic motion. Furthermore, this has also important consequences on the initial minimization of the total energy functional to the ground state, since the number of iterations needed to achieve convergence is proportional to the inverse square root of E_{AG}, and is thus independent of the physical properties of the system (at variance with zero-temperature CP method).

This formulation allows one in principle to simulate efficiently a general physical system, be it a metal or an insulator. There are however some residual problems that affect this method when long molecular dynamics simulations are attempted. On a variety of physical systems we observed a deviation of the electronic and occupational subsystems from the adiabatic regime after a relatively short simulation time. How long one can sustain the simulation depends on the system and the temperature, but for many realistic cases the limitation can be annoying, since one is obliged to periodically quench the electrons to the ground state. A partial solution is to introduce a small friction term in the e.o.m. for electrons and occupations. With an appropriate choice of the friction coefficient this trick works well for structure optimization and annealing, but microcanonical dynamics is hampered by the consequent cooling down of the ions.

Another criticism that can be raised to this scheme is that it is not Lagrangian, i.e. the dynamics for the ξ degrees of freedom is superimposed *ad-hoc* and not derived in a clear way. Moreover, the mass of the ψ variables is time-dependent and such dependency is not taken into account in the equations of motion. Finally, there is no conserved quantity in such formalism. This means that there is no clear way to monitor the energy flux between different subsystems, and therefore it is difficult to analyze the causes which lead to the deterioration of the adiabatic regime.

These issues motivated us to pursue the improvement of the algorythm by combining its efficiency with a well defined mathematical framework. We have been able to show that equations (3.22) and (3.23) can indeed be derived from a Lagrangian function. The exact derivation shows that two velocity-dependent terms that must be included in Eqs. (3.22) and (3.23) for exact energy conservation. The effect of these additional forces on the dynamics will be studied in detail in the following section.

3.5 Lagrangian formulation for metals

In this section we present a fictitious Lagrangian for microcanonical dynamics at finite electronic temperatures. This derivation follows the guidelines of the *Grumbach* scheme, but the *Marzari* occupation matrix formalism is used instead of the simple occupation numbers, as in the *Vandevondele* scheme.

We choose the total free energy of our N-electron system to be the *Marzari* rotationally-invariant A functional:

$$\begin{gathered} A[\{\psi_i\};\{f_{ij}\};\{\mathbf{R}_I\};T] = \\ \textstyle\sum_{ij} f_{ji}\langle\psi_i|\hat{T}+\hat{V}_{ext}[\{\mathbf{R}_I\}]|\psi_j\rangle + \qquad (3.25) \\ +E_{HXC}[n] - TS[\{f_{ij}\}] + E_{MAD}[\{\mathbf{R}_I\}] \end{gathered}$$

where the ψ_i are an orthonormal set of M orbitals, $\mathbf{f}$ is a $M\times M$ positive defined Hermitian matrix with $\mathrm{Tr}\,\mathbf{f} = N$, and the electronic density n is defined as

$$n(\mathbf{r}) = \sum_{ij} f_{ji}\psi_i^*(\mathbf{r})\psi_j(\mathbf{r}) \qquad (3.26)$$

The energy terms which appear in the summation are the electron kinetic energy T, the external potential V_{ext}, the Hartree, exchange and correlation part E_{HXC} and the standard Fermi-Dirac entropy S, which is defined on the eigenvalues f_i of the $\mathbf{f}$ matrix: $S(\mathbf{f}) = \sum_i [f_i \ln f_i + (1-f_i)\ln(1-f_i)]$.

As in *Grumbach*, we don't directly treat the occupations as dynamical variables, but we introduce an auxiliary Hermitian matrix ξ whose elements have the dimension of energy and which can be diagonalized simultaneously with $\mathbf{f}$, via a unitary transformation U: $(U\,\xi\,U^\dagger)_{ij} = \xi_i\delta_{ij}$; $(U\,\mathbf{f}\,U^\dagger)_{ij} = f_i\delta_{ij}$. We define the occupation matrix $\mathbf{f}$ from the eigenvalues of ξ using a Fermi-Dirac function, i.e. $f_i = f_{FD}(\xi_i) = [\exp(\xi_i/T)+1]^{-1}$. We then define the Lagrangian

$$\begin{gathered} \mathcal{L}[\{\psi_i\};\{\dot{\psi}_i\};\{\xi_{ij}\};\{\dot{\xi}_{ij}\};\{\mathbf{R}_\nu\};\{\dot{\mathbf{R}}_\nu\}] = \\ K_f^\psi[\{\dot{\psi}_i\};\{\xi_{ij}\}] + K_f^\xi[\{\xi_{ij}\};\{\dot{\xi}_{ij}\}] + \textstyle\sum_I \frac{1}{2}M_I\dot{\mathbf{R}}_I^2 + \qquad (3.27) \\ -A[\{\psi_i\};\{\xi_{ij}\};\{\mathbf{R}_I\};T] \end{gathered}$$

where the K_f indicate the fictitious kinetic energy contributions and can be written as follows:

$$K_f^{\xi} = -\frac{1}{2} Q \mathrm{Tr}\,(\dot{\mathbf{f}}\,\dot{\xi}), \quad K_f^{\psi} = \mu \sum_{ij} f_{ji} \langle \dot{\psi}_i | \dot{\psi}_j \rangle \tag{3.28}$$

It is easy to see that $\mathcal{L}$ is rotationally invariant [1], thus, as in *Vandevondele*, we choose to write down the Eulero-Lagrange equations of motion in the instantaneous reference frame $\mathcal{O}$ where the ξ matrix is diagonal. In $\mathcal{O}$ the set of coupled differential equations reads:

$$\begin{aligned} (-\bar{f}_{ij} Q)\ddot{\xi}_{ij} &= \bar{f}_{ij}(\xi_{ij} - h_{ij} - \lambda\delta_{ij}) + F_{ij}^{vel-\xi} && (3.29) \\ (f_{ll}\mu)|\ddot{\psi}_l\rangle &= -f_{ll}\hat{H}|\psi_l\rangle + \sum_k \Lambda_{kl}|\psi_k\rangle + |F_l^{vel-\psi}\rangle && (3.30) \\ M_I \ddot{\mathrm{R}}_I &= -\sum_l f_{ll}\langle\psi_l|\frac{\partial \hat{H}}{\partial \mathrm{R}_I}|\psi_l\rangle && (3.31) \end{aligned}$$

where the Lagrange multipliers λ and Λ_{kl} have been introduced in order to enforce charge conservation and orthonormality constraints, respectively. The f_{ll} are the diagonal elements of the occupation matrix, and in $\mathcal{O}$ they correspond to its instantaneous eigenvalues f_l. The quantity $\bar{f}_{ij}$ is defined in $\mathcal{O}$ as

$$\bar{f}_{ij} = \frac{\partial f_{ij}}{\partial \xi_{ij}} = \begin{cases} \frac{f_j - f_i}{\xi_j - \xi_i} & \xi_i \neq \xi_j \\ \frac{\partial f_i}{\partial \xi_i} = g_i & \xi_i = \xi_j \end{cases} \tag{3.32}$$

and is negative definite, since the Fermi distribution function is monotonically decreasing. The diagonal elements $\bar{f}_{ii}$ correspond to the derivative $g_i = g(\xi_i)$ of the distribution function with respect to ξ, while the offdiagonal ones are finite differences in case of nondegenerate ξ. The quantities in brackets, $-\bar{f}_{ij}Q$ and $f_{ll}\mu$, play the role of the masses associated to the ξ_{ij} and $|\psi_l\rangle$ degrees of freedom, respectively. These masses are constructed in order to exactly cancel the prefactors $\bar{f}_{ij}$ and f_{ll} in the main force contributions on the right hand sides, and thus assure an optimal compression of the frequency spectrum of the time evolution of the fictitious degrees of freedom. The drawback is that these masses are no longer constant in time, as they depend explicitly on the ξ_{ij} variables, and as a consequence the velocity-dependent force contributions $F_{ij}^{vel-\xi}$ and $|F_l^{vel-\psi}\rangle$ appear in the EL equations. We stress that, if $F_{ij}^{vel-\xi}$ and $|F_l^{vel-\psi}\rangle$ were absent, the set of equations would be identical to the *Vandevondele* scheme. In the following we will investigate in detail the role played by the additional velocity-dependent terms.

3.5.1 Velocity-dependent forces

The explicit expressions for $F_{ij}^{vel-\xi}$ and $|F_l^{vel-\psi}\rangle$ (we report the details of the derivation in Appendix A) are:

$$F_{ij}^{vel-\xi} = Q \sum_k \mathcal{F}_{ijk}\,\dot{\xi}_{ki}\dot{\xi}_{jk} + \mu \bar{f}_{ij}\langle\dot{\psi}_i|\dot{\psi}_j\rangle \tag{3.33}$$

$$|F_l^{vel-\psi}\rangle \quad = \quad -\mu \sum_k \bar{f}_{kl} \dot{\xi}_{kl} |\dot{\psi}_k\rangle, \tag{3.34}$$

where $\mathcal{F}_{ijk}$ is invariant under any permutation of the three indices, and is defined as:

$$\mathcal{F}_{ijk} = \begin{cases} \frac{\bar{f}_{ij}-\bar{f}_{ik}}{\xi_j-\xi_k} & \xi_j \neq \xi_k \\ \frac{1}{2}\frac{\partial^2 f_i}{\partial \xi_i^2} & \xi_i = \xi_j = \xi_k. \end{cases} \tag{3.35}$$

We can think at both $F_{ij}^{vel-\xi}$ and $|F_l^{vel-\psi}\rangle$ as sums of squared velocities, weighted by the $\mathcal{F}_{ijk}$ and $\bar{f}_{kl}$ prefactors, which depend both on the eigenvalue spectrum of ξ and on the shape of the distribution function $f_{FD}(\epsilon)$. If we suppose a harmonic oscillation regime, $F_{ij}^{vel-\xi}$ and $|F_l^{vel-\psi}\rangle$ are quadratic in the oscillation amplitude, while the leading force terms are linear. However, this is not enough to say that these velocity-dependent forces are small and negligible in a first-order perturbative approach, since this is valid only if the above mentioned prefactors are about of the same order of magnitude as the first-order force constants for a given degree of freedom. Recalling that the main force constants are proportional to $\bar{f}_{ij}$ and f_{ll} respectively for ξ_{ij} and $|\psi_l\rangle$ variables, we introduce the coefficients

$$c_{ij} \quad = \quad \max_k \mathcal{F}_{ijk}/\bar{f}_{ij} \tag{3.36}$$

$$C_l \quad = \quad \max_k \bar{f}_{lk}/f_{ll} \tag{3.37}$$

as an estimate of the relative importance of the velocity-dependent forces in the dynamics of the two subsystems. When these coefficients are large we expect a dynamical evolution dominated by velocity-dependent forces, where the leading forces that keep the electronic degrees of freedom close to the minimum of A may lose effectiveness, and this could induce a significant deviation of the ionic potential from the BO surface. Moreover, in such a regime the electronic or occupational degrees of freedom, whose masses share the same prefactor as the main restoring forces, may experience large and abrupt accelerations, and their trajectory could be difficult to follow for the discrete integrator, or may even diverge.

3.5.2 Numerical example

To understand the general trends in the behaviour of c_{ij} and C_l we calculate these quantities in an ideal case, which is representative of the situation one encounters in realistic cases. We take an evenly spaced eigenvalue spectrum for the ξ matrix, centered at zero and of total width 40, and a smearing energy of unity (arbitrary energy units). We show in Fig. 3.1 as a contour plot the value of $\bar{f}_{ij}$ as a function of ξ_i and ξ_j, and in the right panels we report the shape of the Fermi-Dirac distribution function (upper panel) and of its energy derivative $g(\xi_i)$ (lower panel). We note that $f(\xi_i)$ is vanishing for large positive values of ξ, while $g(\xi_i)$, which is identical to $\bar{f}_{ii}$ (i.e. to the values lying along the dashed

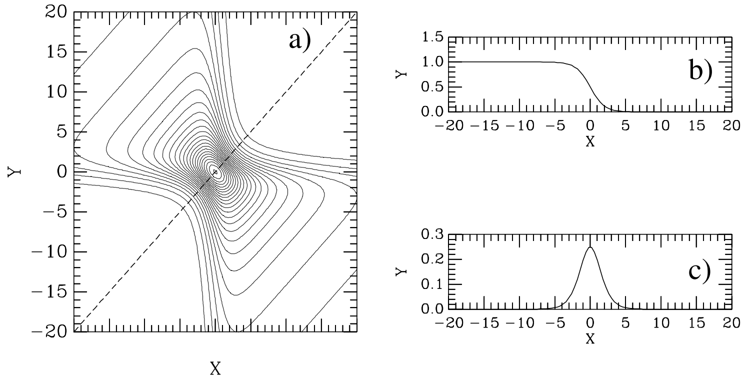


Figure 3.1: (a) contour plot showing the continuous two-variable function $|G(\xi_i, \xi_j)|$ which represents the mass associated to a ξ_{ij} degree of freedom; (b) Fermi-Dirac distribution function $f(\xi_i)$; (c) opposite of the derivative of the distribution function $g(\xi_i)$. A smearing $\sigma = 1$ is adopted, while the eigenvalues ξ_i are evenly spaced between -20 and +20. All energies are in arbitrary units.

line of the figure), is very small at both ends of the spectrum (both functions reach values as small as $\sim 10^{-9}$). Thus, since the vanishing values appear in the denominator of c_{ij} and C_l, we expect large values for the latter quantities in the regions far away from the Fermi energy. The left panels of Fig. 3.2 confirm this expectation. In the upper contour we report the calculated $|c_{ij}|$, which attain values as large as 10^6 (see the logarythmic plot of the diagonal $|c_{ii}|$ values in the middle panel) in the lower left and upper right corners. The C_l coefficients have also an exponentially diverging behaviour, but only in the upper part of the spectrum (lower panel).

This indicates that in realistic simulations, where the smearing energy is much smaller than the total bandwidth (1/40 in our ideal case), at the extreme ends of the eigenvalue spectrum the dynamics of both ψ and ξ becomes strongly dominated by the velocity-dependent forces. In such regime the range of stability of the equations is extremely tight, and the smallest perturbation, due for example to the ionic motion, determines an abrupt upward deviation of the trajectory of the highest eigenvalues ξ_i, a downward diverging trajectory of the lowest ones, or a strong perturbation in the evolution of the unoccupied wavefunctions. All these three cases of instability have been observed in practice, and they eventually caused the failure of the integrator and the collapse of the simulation.

3.5.3 Correction

To overcome this problem we propose a simple solution that allows for stable microcanonical simulations with the Lagrangian equations of motion. We slightly modify the total free energy functional by writing the $f(\epsilon)$ function as a linear

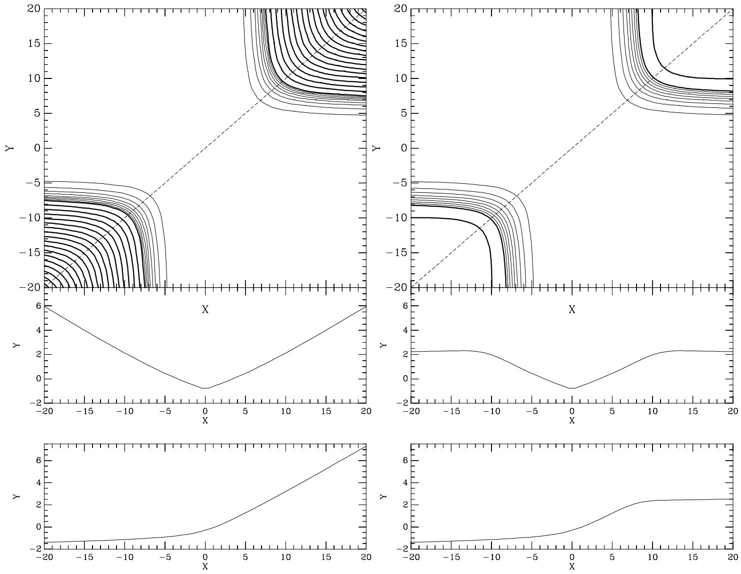


Figure 3.2: Effects of the correction with $\alpha = 10^{-3}$ and $\sigma_2 = 10$ (same units and assumptions as in Fig. 3.1). Left panels show the uncorrected coefficients ($\alpha = 0$) c_{ij} (contour plot in upper panel), the diagonal values c_{ii} (middle panel) and C_i (lower panel). Right panels show the corresponding values when the correction is applied. Notice the different meaning of the contours: thin lines represent equal steps of magnitude unity, thick lines represent logarythmic steps (4 contours=10×). The scale in the xy plots is logarythmic.

combination of two different Fermi-Dirac distributions, the main one with the "correct" smearing energy, and a much broader one with a small prefactor α:

$$f(\epsilon) = (1-\alpha) f_{FD}^{\sigma_1}(\epsilon) + \alpha f_{FD}^{\sigma_2}(\epsilon) \tag{3.38}$$

This modified $f(\epsilon)$ is the exact minimum of a free energy functional where the entropic contribution $\sigma S(f)$ is replaced in the same fashion with $(1-\alpha)\sigma_1 S(f_1)+$ $\alpha\sigma_2 S(f_2)$. To show graphically the effectiveness of this correction, we report in the right panels of Fig. 3.2 the functions calculated with $\alpha = 10^{-3}$ and $\sigma_2 = 10$. The extreme values for c_{ij} and C_l have been quenched to $\sim 10^2$, i.e. four orders of magnitude smaller than the uncorrected quantities. This unphysical feature introduces a systematic error in the resulting ion dynamics. However, in all our tests, a rather small value ($\alpha \sim 10^{-3}$) was sufficient to ensure stability. As a consequence, the average error in the ionic forces due to this modification resulted in all cases to be well below one percent of the total, and thus we consider it negligible in the context of typical CP simulation runs.

3.5.4 Implementation

Despite their apparent complexity, the implementation of the equations of motion in a standard CP electronic structure code is not particularly difficult. Nevertheless, some care must be taken in the correct application of the holonomic constraints and in the evaluation of the velocity-dependent forces. For the wavefunction orthogonalization with mass preconditioning in the standard CP method, Tassone *et al.* [35] proposed the following matrix equation:

$$\mathbf{X}\mathbf{M}\mathbf{X}^\dagger + \mathbf{B}\mathbf{X}^\dagger + \mathbf{X}\mathbf{B}^\dagger = \mathbf{1} - \mathbf{A}, \tag{3.39}$$

Our Lagrangian has an occupation matrix that is no longer proportional to unity, so the corresponding equation in our case reads:

$$\mathbf{F}^{-1}\mathbf{X}\mathbf{M}\mathbf{X}^\dagger\mathbf{F}^{-1} + \mathbf{B}\mathbf{X}^\dagger\mathbf{F}^{-1} + \mathbf{F}^{-1}\mathbf{X}\mathbf{B}^\dagger = \mathbf{1} - \mathbf{A} \tag{3.40}$$

where $\mathbf{F}$ is the occupation matrix, which is diagonal. With a redefinition of the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{M}$ and $\mathbf{X}$:

$$\begin{array}{ll} \tilde{\mathbf{X}} = \mathbf{K}\mathbf{X}\mathbf{K} & \tilde{\mathbf{B}} = \mathbf{L}\mathbf{B}\mathbf{L} \\ \tilde{\mathbf{A}} = \mathbf{L}\mathbf{A}\mathbf{L} & \tilde{\mathbf{M}} = \mathbf{L}\mathbf{M}\mathbf{L}, \end{array} \tag{3.41}$$

where $\mathbf{K} = \mathbf{1}/\sqrt{\mathbf{F}}$ and $\mathbf{L} = \sqrt{\mathbf{F}}$, equation (3.40) can be rewritten in the same form as (3.39) and solved in exactly the same way for the unknown matrix $\tilde{\mathbf{X}}$. Concerning the charge conservation constraint, we note that the chemical potential is defined implicitly in our formalism, since the Fermi level with respect to the eigenvalue spectrum of the ξ matrix is always zero. Therefore, the diagonal elements of the ξ matrix do *not* oscillate about the eigenvalues of the KS Hamiltonian, but are always rigidly displaced by μ. For the computation of the velocity-dependent forces an accurate estimate of $\dot{\psi}(t)$ and $\dot{\xi}(t)$ is needed. The usual Verlet expression $\dot{x}(t) = (x(t+\Delta) - x(t-\Delta))/(2\Delta)$ cannot be directly applied, because the position at $t+\Delta$ is not available and depends on the forces themselves. A fairly accurate solution is the three-point extrapolation

$$\dot{x}(t) = \frac{1}{2\Delta}(3x(t) - 4x(t-\Delta) + x(t-2\Delta)). \tag{3.42}$$

This is sufficient in many cases, but has the undesirable consequence that the integration is not symmetric under time-reversal, and a drift can appear in the evolution of the mathematical constant of motion. Whenever the drift should be too serious, these "predicted" velocities can be improved systematically by taking one or more "corrector" steps [89]. We will show in practice that this technique is able to completely remove the drift.

3.5.5 Tests

We have chosen as test cases the Si_2 dimer, graphitic carbon, and liquid aluminum. These systems are difficult to treat with the standard Car Parrinello [33]

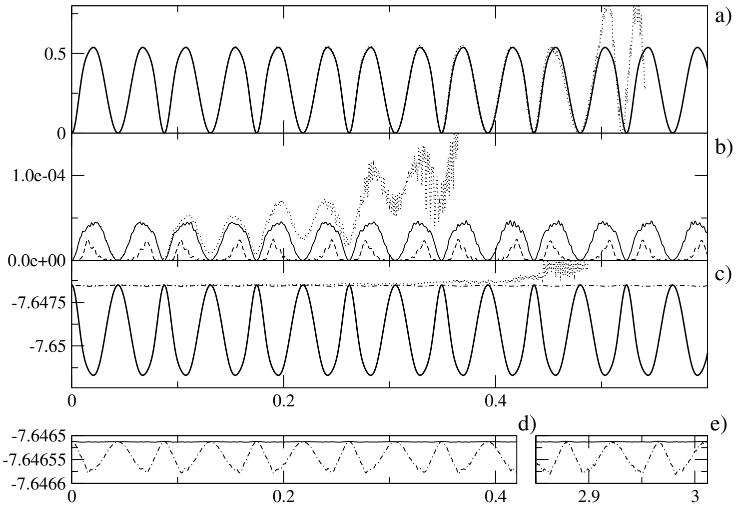


Figure 3.3: Time evolution of various energy cotributions for the Si_2 dimer: (a) ionic temperature (units of 1000 K); (b) K_f^{ψ} (solid line) and K_f^{ξ} (dashed line); (c) potential energy (solid line) and “physical” constant of motion (dot dashed line); (d) blow-up of the physical (dot dashed) and mathematical constant of motion; (e) the same as (d), after ~ 3 ps. of simulation. Dotted lines in the three upper panels show the evolution of the same quantities in an analogous run without velocity-dependent forces.

technique, because Si_2 undergoes a periodic level crossing at the Fermi level during oscillations, and Al is a free-electron metal with a high density of states at E_F. We also present results for graphite, since this system has low-lying electron states located at ~ -20 eV, which provide a stringent test on the stability of our method, which is mainly threatened by eigenvalues lying far away from the Fermi level. In all cases we used LDA for exchange and correlation (spin polarized in the case of Si) and nonlocal norm-conserving pseudopotentials including d orbitals for Si, but only p ones for C and Al.

3.5.6 Si_2 dimer

We initialize the system with a Si-Si distance of 2.032 A, corresponding to ~ 5 % compression with respect to the calculated equilibrium distance (2.152 A). $M = 8$ states were used in the simulation for $N = 8$ electrons, with a cubic cell of 16.0 a.u. and a cutoff energy of 9 Ry. A spin-polarized functional is mandatory, since the electronic ground state of the system has spin $S = 1$, and the level crossings occour between three states of the minority spin manifold, i.e. the two degenerate $3p\pi$ and the singlet $3p\sigma$. We used a timestep $\Delta = 15.0$ a.u., $\mu = 300$ a.u., $Q = 1000$ a.u., a mass preconditioning cutoff [35] of 1.0 Ry,

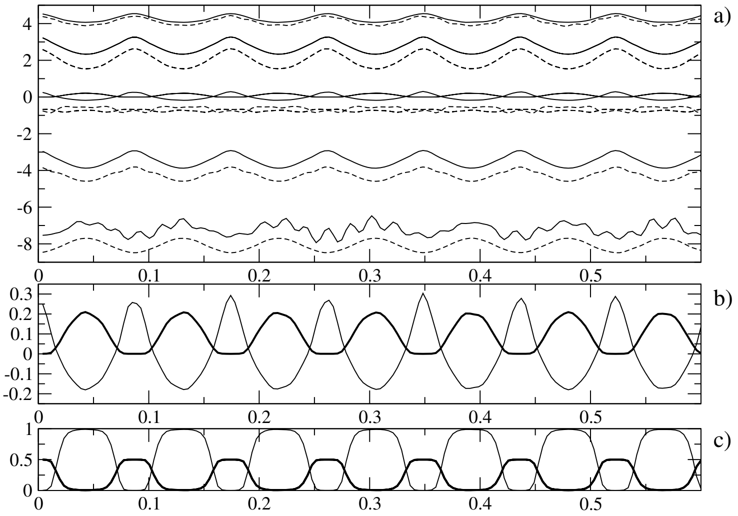


Figure 3.4: (a) time evolution of the eigenvalues of the ξ matrix, corresponding to majority (dashed lines) and minority (solid) spin manifolds; (b) blow-up of the periodic level crossings near E_F, thick line indicates two degenerate π states; (c) occupation numbers of the states depicted in (b).

$\alpha = 10^{-3}$, $\sigma_1 = 0.04$ eV and $\sigma_2 = 6.0$ eV. In Figs. 3.3 and 3.4 we show the results for two microcanonical runs, one employing the full EL equations, and the other suppressing the velocity-dependent forces (i.e. *Vandevondele* scheme). Exactly the same parameters and initial values were used in both cases. The dynamics generated by the full EL equations is very regular, and does not show any sign of deterioration even after ~ 3 ps., i.e. more than 30 full oscillations. On the other hand, the neglect of the velocity-dependent forces produces a fast departure of the electronic degrees of freedom from the BO surface, with a progressive degradation of the ionic trajectories, which become unacceptable after four or five oscillations. Fig. 3.4 shows the evolution of the ξ_i eigenvalues, evidencing the level crossings near the Fermi level. The π multiplet starts with eigenvalue approximately zero, corresponding to half filling, and its occupation number decreases to virtually zero as the ions go away from each other. The occupation of the σ state raises from zero to one as the multiplet loses importance. We notice that for a stable treatement of the periodic filling of the states at E_F the smearing energy σ_1 has to be chosen carefully, because values that are too low can lead to abrupt changes in the charge density and self-consistent potential, with substantial difficulties for the electronic degrees of freedom to evolve abiabatically. In particular, for similar tests with higher oscillation amplitudes we had to choose between two alternatives, i.e. increasing σ_1 or reducing the fictitious masses and the integration timestep according to the higher ion ve-

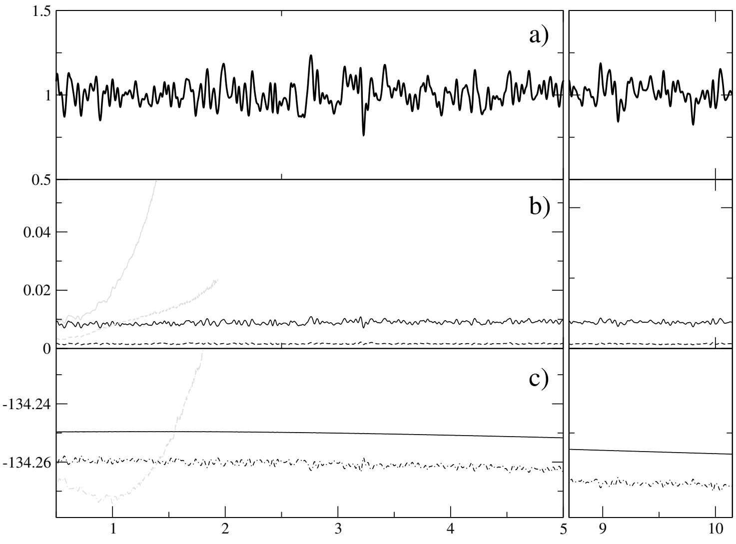


Figure 3.5: Time evolution of various energy cotributions for the liquid Al system: (a) ionic temperature (units of 1000 K); (b) K_f^{ψ} (solid line) and K_f^{ξ} (dashed line); (c) physical (dot dashed) and mathematical (solid) constants of motion. Right panels: after $\sim$ 10 picoseconds. Both c.o.m. in (c) have a drift of approximately 0.4 meV/(at.ps). Grey lines show the evolution of the same quantities in an analogous run without velocity-dependent forces.

locities during level crossings. On the other hand, the use of a larger smearing allows one to use larger fictitious masses and longer timesteps.

3.5.7 Liquid Al

For the test on the liquid Al system we set up a cubic supercell of 20.65 a.u. containing 64 Al atoms whose positions were generated by a previous microcanonical run at $\sim$ 1000 K. We represented the electronic structure on $M = 256$ states with an electronic temperature $\sigma_1 = 0.75$ eV. Other parameters were assigned the following values: $\Delta = 40.0$ a.u., $\mu = 3200$ a.u., $Q = 9600$ a.u., a mass preconditioning cutoff of 1.0 Ry, $\alpha = 10^{-3}$ and $\sigma_2 = 6.5$ eV. We started the simulation with zero ionic velocities, we accelerated the ions during the first 500 steps, and then we left the system free to evolve for $\sim$ 10 ps. at an average temperature of 1000 K. We report in Fig. 3.5 the evolution of various energy terms during the simulation. The physical constant of motion $W_{phys} = K_{ion} + A$ is almost flat, and has a drift as low as $\sim$ 0.4 meV/(at.ps) thus indicating a high quality adiabatic behaviour that is preserved even after such a long simulation time. Also in this case the contribution of the velocity-dependent forces is quite important, as one can see from the comparison with an analogous run based

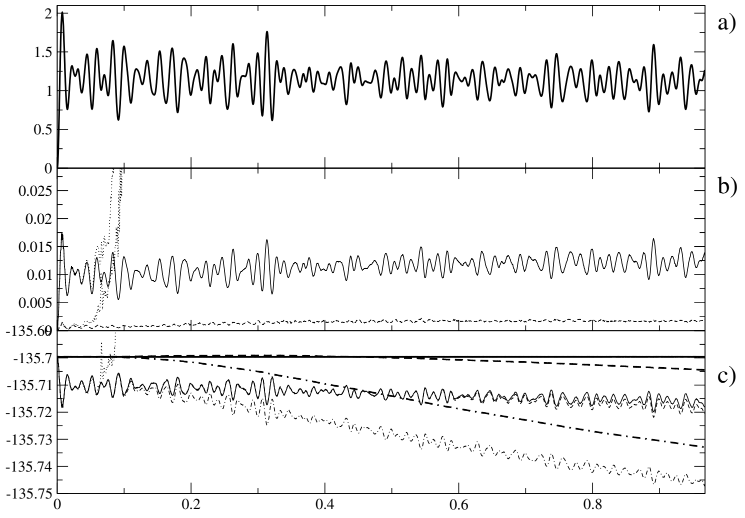


Figure 3.6: Time evolution of various energy cotributions for the graphite system: (a) ionic temperature (units of 1000 K); (b) K_f^{ψ} (solid line) and K_f^{ξ} (dashed line); (c) comparison of the physical (thin lines) and mathematical constant of motion (thick lines) in three runs with uncorrected velocities (dot dashed), with one corrector step (dashed) and with five corrector steps (solid lines). Dotted lines show the evolution of the same quantities in an analogous run without velocity-dependent forces.

on the *Vandevondele* equations. We notice that the optimal performance of our method allowed for a perfectly stable simulation with an unusually large integration time step ($\sim$ 1 fs). After 10 ps, the gain of electronic kinetic energy is negligible, and the drift in W_{phys} is dominated by error accumulation in the integration of the velocity-dependent terms. In this test we used only one corrector step, that is sufficient to ensure a very small drift of W_{phys}. More corrector steps per iteration would result in an even smaller drift.

3.5.8 Graphite

For the graphite system we set up a supercell containing 24 carbon atoms in one layer, with a nearest neighbor distance of 1.45 Å, and an interlayer distance of 3 Å. We represented the electronic structure on $M = 72$ states with an electronic temperature $\sigma_1 = 0.5$ eV. The other parameters of the simulation are $\Delta t = 8.0$ a.u., $\mu = 300$ a.u., $Q = 1000$ a.u., $E_C^{mass} = 3.0$ Ry, $\alpha = 2 * 10^{-3}$ and $\sigma_2 = 13$ eV. The starting atomic coordinates are generated by a random displacement from the equilibrium positions, and starting velocities are set equal to zero. Then the system is left free to move for $\sim$ 1 ps, and the ions oscillate at a temperature slightly above 1000 K. Inspection shows that the system is metallic, since there is a multiplet of 4 states at the Fermi level sharing 4 electrons. These states are degenerate at equilibrium, and during the simulation at 1000 K they all periodically cross the Fermi level. Using the same initial conditions and simulation parameters we tested the performance of the integration algorythm by doing three different runs: one in which the velocities were calculated with expression (3.42), a second run in which one corrector step was added, and a third run with five corrector steps. The results for the evolution of the mathematical (W_{math}) and physical constants of motion in the three cases are given in Fig. 3.6. Without corrections the evolution is numerically well behaved, but W_{phys} shows an important drift, corresponding to 45 meV/(at.ps), and W_{math} has the same behaviour, indicating that the extrapolated velocities are not sufficiently accurate in this case. The second run with only one corrector step shows an important improvement in the evolution of both quantities, with a drift that is now about 12 meV/(at.ps) for W_{phys}. The third run has the velocity-dependent forces well converged and consistent with the trajectories to machine precision. As a consequence, W_{math} is stable. However, the evolution of W_{phys} was similar to the previous case, with a drift of 7 meV/(at.ps) due to the energy transfer from the ionic to the electronic subsystem. For comparison, we report in the same figure as thin dotted lines the behaviour of some of these quantities, in a similar run where the velocity forces have been neglected. In this test the electronic and occupational variables depart quickly from the BO surface, thus confirming the importance of accurate velocity-dependent terms in ensuring a metastable adiabatic regime. We notice that this system is relatively "difficult" to treat using our finite-temperature Lagrangian, and that in less severe conditions the integration without corrections should be sufficient.

3.5.9 Discussion and outlook

We have proposed a fictitious Lagrangian for *ab-initio* molecular dynamics simulations of metallic systems at finite electronic temperatures. This formulation extends the standard Car-Parrinello method to systems with variable and fractional occupations, and shows substantial improvements with respect to previous techniques in some particular cases. The correct treatment of the Lagrangian degrees of freedom ensures energy conservation, which is highly desirable, but has a serious drawback in that the equations of motion are no longer linear, and instabilities can occour in the limit of low electronic temperatures. In order to circumvent these instabilities, we had to modify the Fermi distribution function in an unphysical and somewhat disturbing way, i.e. we had to assign a finite occupancy to the empty states at the expense of the low-lying valence states. This procedure, apart from the introduction of systematic errors in all physically relevant quantities (which we have shown to be relatively small in our examples) produces negative unwanted consequences:

- the overall convergence to the ground state of the initial wavefunctions is dramatically affected;
- the unavoidable "heating up" of the high unoccupied states during the

simulation is enhanced by their finite occupation number.

This second feature is clearly evident in our tests on graphite. The drift in the *physical* constant of motion, which was important in this case, could not be reduced simply by introducing more empty states in the simulation, at variance with the *Vandevondele* method. These shortcomings limit the *general* applicability of our Lagrangian formulation, and is most serious when localized semicore states are used, e.g. in transition metals. However, this study was extremely valuable in identifying the actual problems and in suggesting a strategy to overcome them.

The most serious problem affecting the *Vandevondele* scheme is found to be the progressive accumulation of kinetic energy in the highest unoccupied empty states. When the velocities of these states start to be important, error accumulation due to the inexact treatment of the varying masses causes a rapid departure of the ψ and ξ subsystems from the adiabatic regime, and a periodic "quench" to the ground state is needed in order to extend the simulation over long time scales. This behaviour is largely corrected by the time-dependent Lagrangian forces, but the price to pay is the instability of the resulting equations which seriously hampers the robustness of the formulation.

We notice that these problems occour only in microcanonical simulations, while in electronic minimizations to the ground state, structure relaxations or simulated annealing the Vandevondele scheme performs optimally and has been successfully used in a number of studies [3, 12, 5]. In fact, during structure relaxations, a friction term on the fictitious degrees of freedom avoids any progressive accumulation of kinetic energy in the unoccupied states, and prevents the electronic subsystem from "exploding". Thus, a small friction term could allow for fairly accurate simulations of chemical reactions and phenomena that need a relatively short simulation time, but longer microcanonical runs would still be difficult because of the progressive cooling down of the ionic subsystem.

For future improvement, one could e.g. introduce better friction terms than the viscous one, in order to damp out effectively the high empty states only, which are negligibly coupled to the ions and thus avoid any appreciable decrease of the average physical temperature. Such a strategy has not been tried yet, but seems promising since the role of the highest unoccupied states is easily tuned (it is a computational parameter which depends on the electronic smearing temperature and on the number of empty wavefunctions included in the calculation). In this way one renounces to the formally elegant energy-conserving Lagrangian framework, but in actual simulations what matters is the smooth behaviour of the physical quantities, and a robust and fully controllable algorythm for the electronic subsystem. Nonetheless, our derivation helped to corroborate the mathematical soundness of the Vandevondele approach, to understand the residual problems, and to find out possible ways to circumvent them.

Chapter 4

C_{60}/Al(111)

In this chapter we present a first-principles study of the adsorption of C_{60} molecules on the Al(111) surface, with particular attention to the Al-C bonding character and its role in determining the equilibrium atomic configurations. Apart from the fundamental and technological interest in the detailed nature of C_{60}/Al bonding, this study allows for a practical test of the novel algorythms presented in the previous chapter. Large-scale, metallic, highly inhomogeneous systems are difficult to treat by means of *ab-initio* approaches, and the availability of a stable and efficient technique was a key factor for the success of the present study.

4.1 C_{60}: structural and electronic properties

C_{60} has the shape of a soccer ball, consisting of a bonding pattern of 12 pentagons and 20 hexagons, with carbon atoms occupying the vertices (see Fig. 4.1). It can be thought as a graphite sheet folded onto itself, with the 12 pentagons satisfying the topological requirements for the obtention of a closed polyhedron. The two inequivalent C-C bond lengths have been experimentally measured with high accuracy [39] and the reported values are 1.40 and 1.46 Å, for the edges linking a pentagon (5-ring) and an hexagon (6-ring) and those linking two hexagons, respectively. These values can be compared with 1.39 Å for sp^2 graphite and 1.54 Å for sp^3 diamond, and indicate for C_{60} a predominant sp^2 bonding picture, slightly perturbed by angular distortions.

C_{60} molecules interact among themselves through relatively weak van der Waals forces, which are similar in nature and strength to graphite interlayer interactions [40]. They arrange into a molecular FCC solid called fullerite with a nearest neighbour distance (NND) of 10.04 Å at room temperature. This solid phase displays a considerable orientational disorder since the individual molecules are free to rotate above 260 K, thus indicating a rather weak anisotropy for the C_{60}-C_{60} interactions [1] Fullerite sublimates at 450-500 K, and

---

[1] At low temperature the rotational motion is frozen, and the stable structure is cubic with

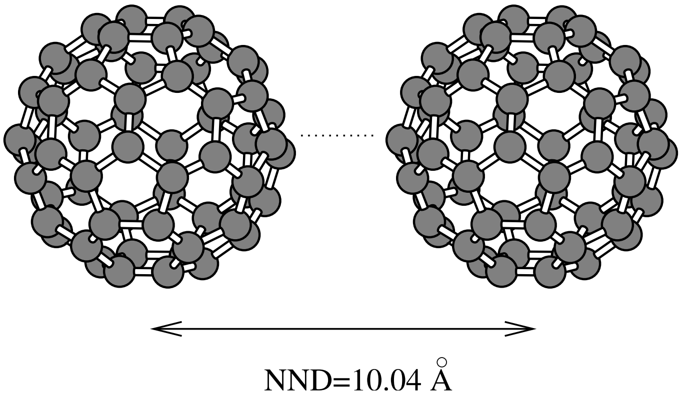


Figure 4.1: Schematic representation of the marked difference between intermolecular and intramolecular interactions of C_{60} molecules: strong covalent bonds lead to stiff icosahedral cages which attract each other through weak van der Waals forces.

its measured cohesive energy is $E_{coh} = 1.739$ eV/molecule [41].

From the value of E_{coh} we can directly estimate the energy of a single C_{60}-C_{60} bond, which is simply $E_{coh}/6 \sim 0.3$ eV in an independent pairwise forces approximation. Now we can estimate the adsorption energy of an isolated C_{60} molecule at the (111) surface of solid fullerite, which is $E_{ads} \sim 0.9$ eV (the adsorbed C_{60} is in contact with three other C_{60} of the substrate). This estimate is quite useful, since it sets a lower bound for E_{ads} on substrates where higher desorption temperatures are reported, in particular Al(111). In the following section we will give a general overview about what is known on the adsorption of C_{60} on different crystalline surfaces, and then we will introduce the particular features of C_{60}/Al(111) which are known from eperimental data.

4.2 C_{60} adsorption on surfaces

On a perfectly flat and homogeneous surface, C_{60} molecules would be free to move along the plane, and would arrange themselves in a perfect hexagonal layer with the NND of solid fullerite. Real crystal surfaces, even in the ideal case, have obviously a certain degree of corrugation, which leads to a more or less marked site preference for C_{60} adsorption, and to energy barriers for in-plane diffusion at low temperature. In Fig. 4.2 we show schematically the role of these competing contributions, i.e. C_{60}-C_{60} and C_{60}-substrate interactions in determining the equilibrium structure for a well annealed C_{60} overlyer.

Detailed information for adsorption geometries, nature of bonding and elec-

---

four C_{60} molecules per unit cell.

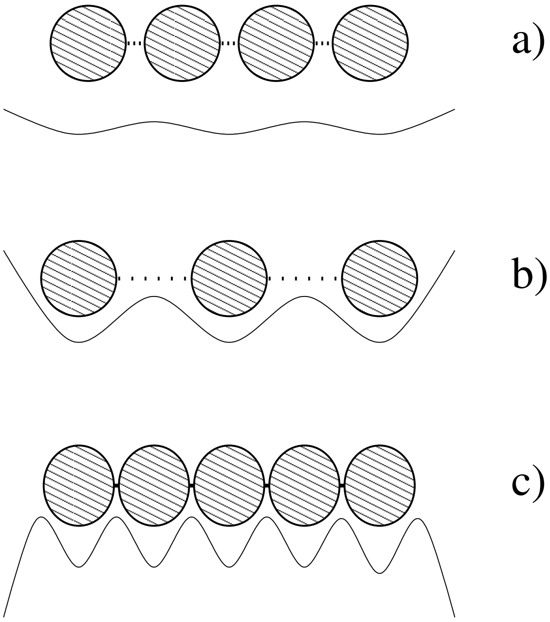


Figure 4.2: Adsorbate-adsorbate and adsorbate-substrate interactions play competing roles in determining the equilibrium overlayer structure. When the corrugation of the in-plane potential energy surface is weak, C_{60}-C_{60} interactions dominate and uncommensurate structures with the NND of solid fullerite are formed (a); when the interaction with the substrate is strong, the overlayer is forced into commensurate structures, with some tensile (b) or compressive (c) strain in the overlayer in order to match the substrate periodicity.

tronic structure have been collected for C_{60} deposited on a wide range of substrates. It is difficult in the context of surface physics to characterize the adsorbate/substrate bonding character on the basis of rigorous and well defined concepts, but following the classification of [55] we can nevertheless rationalize the C_{60} adsorption on different substrates into three main groups: *van der Waals*, *ionic*, and *covalent* bonding. We will briefly describe the characteristic features of these categories by giving representative examples (see e.g. [46] for comprehensive references to specific works).

Van der Waals. On inert surfaces such as GeS(100), graphite, and SiO_2 the desorption temperatures T_{des} for submonolayers of C_{60} are close to the sublimation temperature of solid fullerite, thus indicating interactions of comparable strength, i.e. essentially van der Waals in nature. This picture is confirmed by the observation of nearly ideal hexagonal C_{60} overlayers on these substrates, which suggests that the C_{60}-C_{60} interaction is the dominant factor in determining the equilibrium structure. Adsorbed C_{60} molecules are in general highly mobile at room temperature (RT).

Ionic. C_{60} monolayers on low-index surfaces of noble metals have been extensively studied, and significantly stronger adsorbate/substrate interactions have been found than in previous cases, with T_{des} values close to 800 K. Characteristic of this category is a significant charge transfer from the metal to the C_{60} lowest unoccupied molecular orbitals (LUMOs), which was detected by measuring the shifts in the vibrational modes of adsorbed fullerenes, by observing the position of the LUMO-derived peak in photoemission spectra, or the intensity of the corresponding resonance in x-ray adsorption. STM and x-ray diffraction has been further used to deduce equilibrium structures, mobility and orientational order of adsorbed C_{60}, and from these investigations an overall picture emerges, of predominantly ionic bonds, i.e. charged spheres bound by the image potential at the metal surface.

Nevertheless, as a consequence of the enhanced strength of this interaction compared to inert surfaces, some degree of commensurability with the substrate lattice constant is in general observed. Moreover, upon annealing, well ordered overlayers are obtained which sometimes display puzzling features in STM images, e.g. inequivalent adsorption sites with different apparent heights and electronic properties. These observations can hardly be explained without accepting that some rearrangement of the substrate occours at the atomistic level, and thus they give evidence for overlayer-induced substrate reconstructions. The details of the underlying mechanisms are in general not clear, although some structural models have been inferred from experimental data. Despite the dominant ionic bonding, a realistic cohesive model should account for the hybridization of the C_{60} molecular orbitals with metal states, which was found to be important e.g. in the C_{60}/Au(110) system, and which could play an important role in determining the observed reconstruction patterns.

Covalent. Several semiconductor (Si, Ge) and metal (Pt, Ni) substrates interact much more strongly with C_{60}, forming chemical bonds of predominantly covalent character. The substrate-adsorbate interaction now dominates over intermolecular forces, and adsorbed C_{60} molecules form commensurate structures. As a consequence the C_{60}-C_{60} equilibrium spacings can deviate substantially from the fullerite NND, e.g. 5 % compression and 43 % expansion are reported for Ge(100) and Ge(111) respectively. The adsorbates are in general immobile at RT, and relatively high annealing temperatures are necessary to obtain well ordered overlayers. In many cases the bonds with the substrate are strong enough to catalyze the decomposition of C_{60} prior to desorption, e.g. at a temperature of $\sim$ 1050 K on Pt(111).

4.3 C_{60} on Al(111)

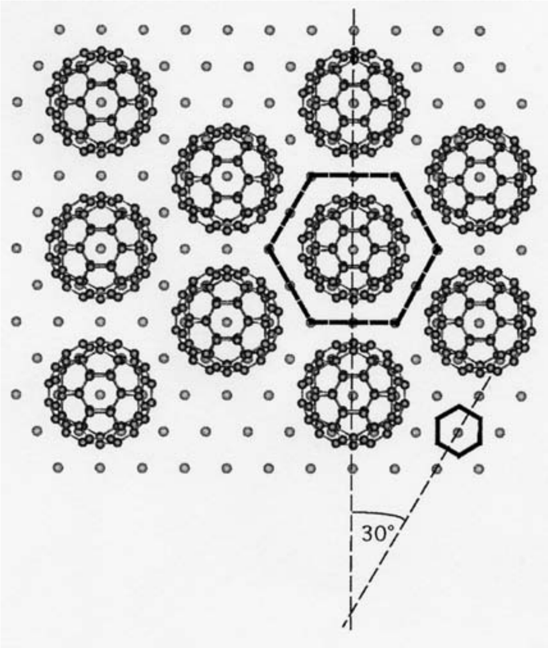


Figure 4.3: A model for the $2\sqrt{3} \times 2\sqrt{3} - R30°$ phase of the C_{60}/Al(111) system. An hexagonal surface unit cell is indicated (dashed hexagon), and is compared to the unit cell of the clean surface (solid hexagon). Notice the 30 degrees rotation, and the angular orientation of the molecules with respect to the substrate. The registry is arbitrarily fixed to an atop position (figure taken from Ref. [46]).

The adsorption of C_{60} on the Al(111) surface displays peculiar and interesting characteristics that deserve separate treatment, and can be summarized as follows:

- the bonding is of covalent character without evidence of charge transfer;
- the desorption temperature is $\sim$ 730 K;
- the molecules are mobile at RT and migrate to step edges at low coverage;

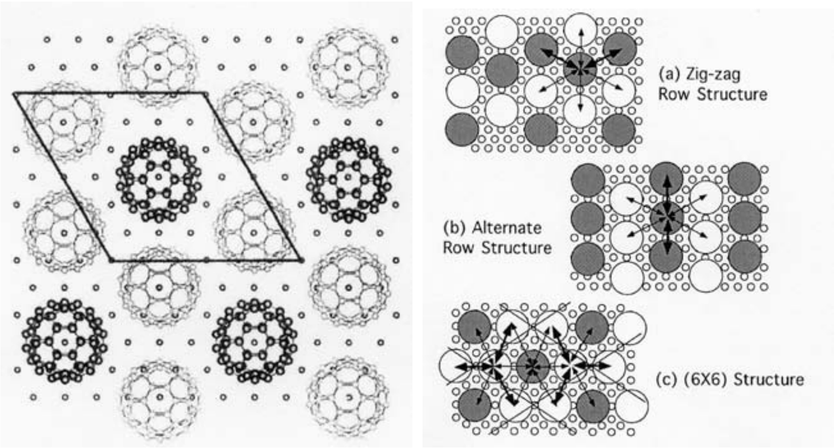


Figure 4.4: Left panel: model for the reconstructed geometry. All C_{60} molecules have a one-to-one correspondence with those in Fig. 4.3, but one out of three protrudes in STM topographs (indicated in darker colour). Right panel: this (6×6) reconstruction pattern is the most efficient way to relieve the in-plane compressional strain while keeping a maximum of C_{60} molecules in contact with the metal (taken from Ref. [46]).

- the overlayer grows commensurate to the substrate with slight in-plane compression (~ 1 % with respect to NND in solid fullerite).

These data indicate that the bonding between C_{60} and Al is not fully compatible with any of the three pictures discussed above, and should be considered representative of a new *intermediate covalent* bonding category. Since this system is the subject of the present theoretical study we will now discuss the experimental results in deeper detail.

For C_{60}/Al(111), two different overlayer structures have been reported. By evaporating C_{60} onto a sample held at room temperature the molecules arrange in an hexagonal overlayer with $2\sqrt{3} \times 2\sqrt{3} - R30°$ symmetry, i.e. 12 Al surface atoms per surface unit cell (a model is shown in Fig. 4.3). Accurate STM topographs were able to resolve the electronic orbitals corresponding to HOMO and LUMO, showing that the molecules are preferentially adsorbed with a 6-ring facing the substrate, the bonds of the lowermost hexagonal ring being parallel to the $[10\bar{1}]$ directions of the Al(111) surface. These results are in good agreement with previous x-ray diffraction studies, but the adsorption site was not obtained in any of these works, so the details of the bonding configuration remain unclear.

Upon annealing at 490 K, a reconstruction is observed where one out of three C_{60} molecules ends up at a higher position, and the raised molecules arrange in a regular (6×6) lattice with one-to-one correspondance with the $2\sqrt{3} \times 2\sqrt{3} - R30°$ sites (see Fig. 4.4). It was noticed that this geometry is the most effective way to relieve the in-plane compressional strain while keeping the maximum number of C_{60} molecules as close as possible to the substrate (see the right panel of Fig. 4.4).

The driving force for the reconstruction cannot, however, stem from C_{60}-C_{60} bond compression alone. Displacing molecules away from the substrate implies breaking a number of C-Al bonds which are much stronger than the intermolecular interactions, consistent with the the observation that C_{60} desorption from Al does not occour below 700 K, while fullerite sublimates at 450-500 K. There exist, meanwhile, experimental indications that the raised molecules may be strongly bound to the substrate, possibly even *more* than the lower ones [53]. These observations could be reconciled if some Al atoms were displaced from their original positions to form bonds with the raised molecules, a picture supported by experimental evidence of Al adatoms in the (6×6) structure [52]. It is, however, unclear where these adatoms may come from. A possibility is that they detach from surface step edges and migrate to their stable location underneath the C_{60} molecules, consistent with the observation that the shape of the steps can change during reconstruction [53]. However, long range mass transport would imply a strong correlation between the step density and the kinetics of the phase transition, and no such dependence has been reported so far. A simpler explanation is that Al adatoms are produced *locally* on C_{60}-covered terraces. However, the high energy cost of forming adatom/vacancy pairs on a clean Al surface (1.7 eV [55]) seems to rule out this mechanism at the reconstruction temperature of 490K.

In this work, we provide evidence that the covalent character of the C_{60}/substrate interaction promotes the reconstruction by the adatom-vacancy mechanism. Namely, we show that the presence of the C_{60} overlayer drastically reduces the energy cost of adatom-vacancy pair creation in the metal substrate, because of an unusual steric effect associated to the C_{60} adsorption bonding geometry on the (111) metal surface. This can yield a reconstruction mechanism where Al vacancies are created at the adsorption sites together with adatoms which move to the interstitial regions underneath the C_{60} overlayer. Large-scale first-principles structure relaxations performed directly on the (6×6) system lead to a reconstructed structure which is significantly more stable than the unreconstructed one and whose unit cell contains two C_{60} molecules bound to vacancies and one Al ad-dimer placed underneath a third C_{60} molecule.

4.4 Computational details

Our methods, which are based on Density Functional Theory within the Local Density Approximation (LDA) for exchange and correlation, describe correctly the strong Al-Al and C-Al bonding [57]. The aluminium surface is modeled by 5-layer slabs separated by a minimum 8 Å vacuum thickness, using surface supercells with $(2\sqrt{3} \times 2\sqrt{3})R30^\circ$ and (6×6) periodicity. The primitive translation vectors in the xy plane correspond to the calculated Al lattice constant of 3.97 Å, which is 2 % less than the experimental value of 4.05 Å. Norm-conserving pseudopotentials [22] are used for the electron-ion interactions, and

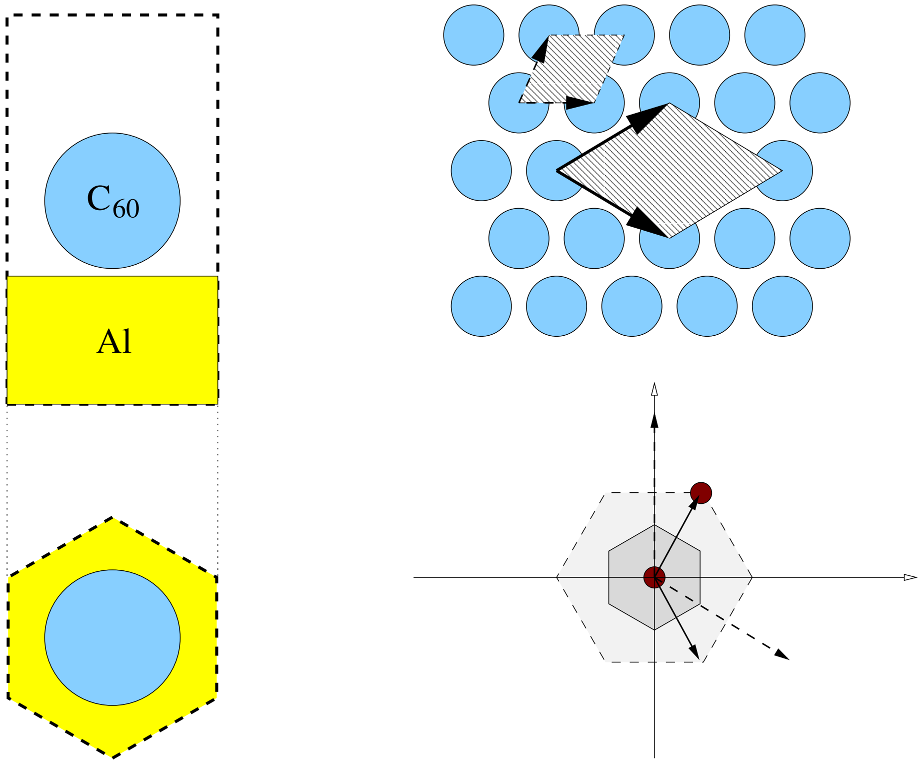


Figure 4.5: Summary of the computational geometric parameters. Left: hexagonal $2\sqrt{3} \times 2\sqrt{3} - R30^\circ$ unit cell used in the study of the unreconstructed configurations. Right: real (top) and reciprocal (bottom) space representation of the surface unit cell and primitive translation vectors. The surface Brillouin zones (BZ) are shown as shaded polygons, the larger corresponding to the $2\sqrt{3} \times 2\sqrt{3} - R30^\circ$ cell. The k-points used for BZ sampling are indicated by red circles. Notice that for mathematically equivalent datasets two k-points are used in the $2\sqrt{3} \times 2\sqrt{3} - R30^\circ$ geometry (all vertices of BZ boundary are equivalent), and only the Γ point in the 6×6 supercell.

Conf.	E$_{ads}$	C-Al	C-C(1)	C-C(2)	C-C(3)	C-C(4)
atop	0.98	2.53	1.44	1.47	1.44	1.47
bridge	1.20	2.20	1.39	1.48	1.45	1.42
hollow-6	1.37	2.24	1.52	1.48	1.39	1.46
hollow-5	1.16	2.28	1.41	1.44	1.45	1.53

Table 4.1: Adsorption energy E$_{ads}$ (eV) and bond lengths (Å)for several unreconstructed structures in the $2\sqrt{3}\times 2\sqrt{3}R30°$ supercell. The adsorption energies are calculated as $E_{ads} = E_{C_{60}} + E_{Al(111)} - E_{C_{60}/Al(111)}$.

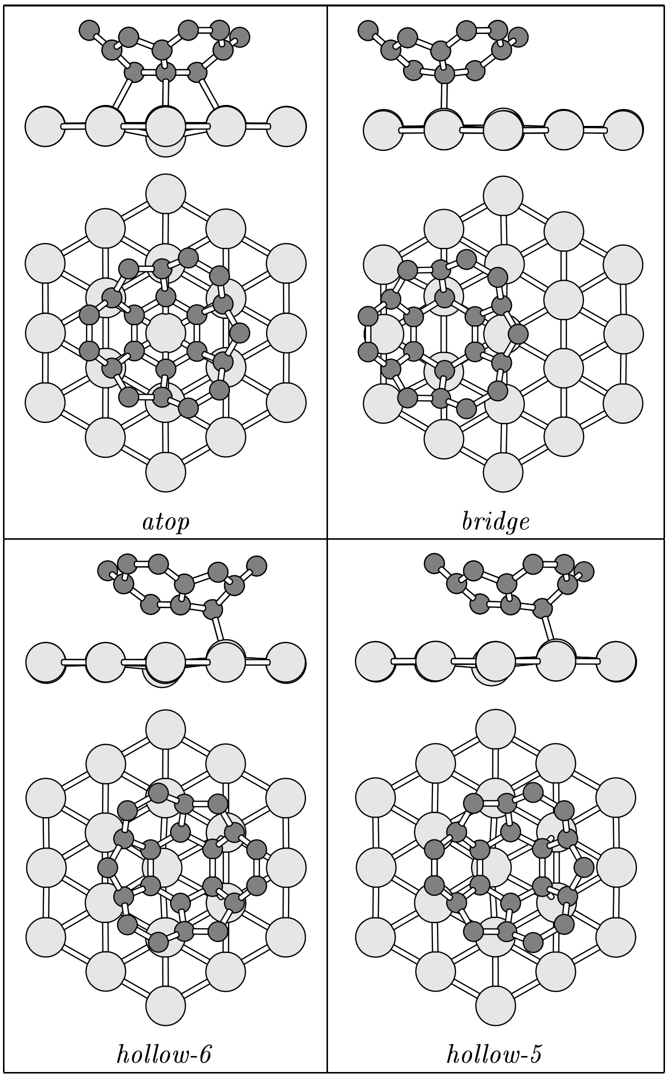
have been generated with the fhi98PP [56] code. A plane-wave cutoff of 35 Ry provides satisfactory convergence of the atomic structures and energy differences; tests performed with a 50 Ry cutoff showed variations of the order of 0.01 eV. Brillouin-zone sampling is performed with the Γ point only for the (6×6) phase, while we use two k-points for the smaller cell in order to have equivalent data sets in the two cases (see Fig. 4.5). We use a finite temperature energy functional with a Fermi-Dirac distribution and a smearing energy of 0.25 eV. All electronic structure minimizations and ionic relaxations are made with the *Vandevondele* scheme, which was extensively discussed in Chapter 3.

4.5 Results: unreconstructed configuration

To study the unreconstructed $(2\sqrt{3} \times 2\sqrt{3})R30°$ system, we orient the C$_{60}$ molecules as experimentally observed [51, 54], i.e. facing the substrate with an hexagonal ring of C-C bonds oriented along $\langle 10\bar{1}\rangle$ directions. Since no information is available on the adsorption site, we investigated several positions in the xy surface plane. The most favorable structure is obtained when two C neighbours belonging to adjacent 6-rings bind to a pair of Al atoms of the first layer (Figure 4.6, *hollow-6* structure) [2]. A similar configuration involving a 5-/6-ring C pair (*hollow-5*) results in a slightly higher total energy (cf. Table 1). When two *opposite* C atoms of the lowest 6-ring are bound to the Al dimer we obtain a third configuration (*bridge*) with a total energy intermediate between the two *hollow* ones. In the position of highest symmetry the molecule is centered over a surface Al atom (*atop*) and six Al-C bonds are formed. The latter geometry, previously speculated to be the stable one [51], corresponds in our calculations to a saddle point and not to a local energy minimum. Relaxing symmetry constraints, the molecule spontaneously shifts along the $\langle 2\bar{1}2\rangle$ direction, thus strengthening two of the six bonds while breaking the other four, and eventually reaches one of the two adjacent *hollow* sites with substantially lower energy. Our adsorption energies can be compared with those of other

---

[2] Each geometry in Fig. 4.6 is associated to two inequivalent adsorption sites, which differ by the stacking of the underlying FCC aluminium (ABC or ACB). A test on the *hollow-6* configuration revealed that stacking differences yield negligible differences in the total energies.



atop

bridge

hollow-6

hollow-5

Figure 4.6: The four geometries considered for the adsorption of C_{60} on the clean Al(111) surface. For clarity, only atoms from the outermost Al layer (light grey) and the lower part of the adsorbed C_{60} molecule (dark grey) are shown.

systems involving carbon bonding to Al(111). For ethylene [60], physisorption is predicted (d_{C-Al}=3.5 Å, E_{ads}=0.03 eV). For benzene, a weak electrostatic interaction (d_{C-Al}=3.7 Å, E_{ads}=0.35 eV) has been reported [61], similarly to thiophene [62]. On the other hand, the CH_3 radical binds strongly to Al(111) (d_{C-Al}=2.0 Å, E_{ads}=1.80 eV/1.96 eV) [63]. Our results for C_{60} are intermediate between those of C_6H_6 and CH_3.

The relaxed atomic configurations show many details that give interesting insight into the nature of C_{60}-Al bonding. The adsorbed molecule is distorted by the interaction with the metal, i.e. the C-C bonds belonging to the lowermost 6-ring are considerably stretched while the upper part of the molecule is practically unperturbed. Concerning the substrate, in the atop configuration the Al atom directly below the C_{60} undergoes a large downward shift (about 0.5 Å, Fig. 4.6), while in all other configurations the substrate relaxation is practically negligible. While the bond stretching values of the adsorbed C_{60} (cf. Table 1) are consistent with known experimental data [46], the unusually large displacement of the *atop* Al atom points to an interesting new effect.

To investigate the electronic structure changes due to the C_{60}/surface interaction we compute the interaction density displacement $\Delta\rho = \rho_{C_{60}/Al(111)} - \rho_{C_{60}} - \rho_{Al(111)}$, using the atomic positions of the C_{60}/Al(111) system for the unrelaxed *atop* configuration. The yz cutplane contour plot in Fig. 4.8 and the isosurfaces in Fig. 4.7 show a strong accumulation of electron density along the C-Al bonds, indicating a partial sp^3 hybridization at the C atoms interacting with the substrate, consistent with the covalent picture previously reported [52, 46]. Interestingly, there is also an electron depletion in the substrate, in particular directly above the central atom of the outermost atomic layer (we will call this the *central* region in the following discussion). This can be explained considering that in the central region the electronic states of the metal surface are stabilized by having components over the atomic orbitals of the central Al atom. This atom, however, does not bind to C atoms, while the formation of C-Al bonds between its six Al neighbors and C_{60} makes hybridization in the central region (in particular, with the p_z orbital of the central Al atom) less favorable due to the orthogonality constraints and to electrostatic repulsion. In the language of VSEPR [58] theory, the newly formed C-Al bonds “repel” the filled orbitals of the central Al atom. This atom is thus expected to move away from the central region if allowed to relax, in order to lower the energy of its filled orbitals. Consistently, during the *atop* geometry relaxation, this atom moves downwards from its ideal position as the C_{60}-Al distance decreases. Eventually an equilibrium is reached due to the repulsion between the “sinking” atom and the subsurface Al atoms. This prevents the C_{60} from approaching the surface as close as the optimal Al-C bond length would allow (cf. Table I), explaining why the bonding in the *atop* configuration is weaker than in the other three geometries, although more C-Al bonds are formed.

This “uncomfortable” position of the central atom is a key result, as it suggests a possible mechanism for overlayer-induced adatom-vacancy pair production. We will discuss this point in the following section, and then we will show how this could play an important role in the experimentally observed

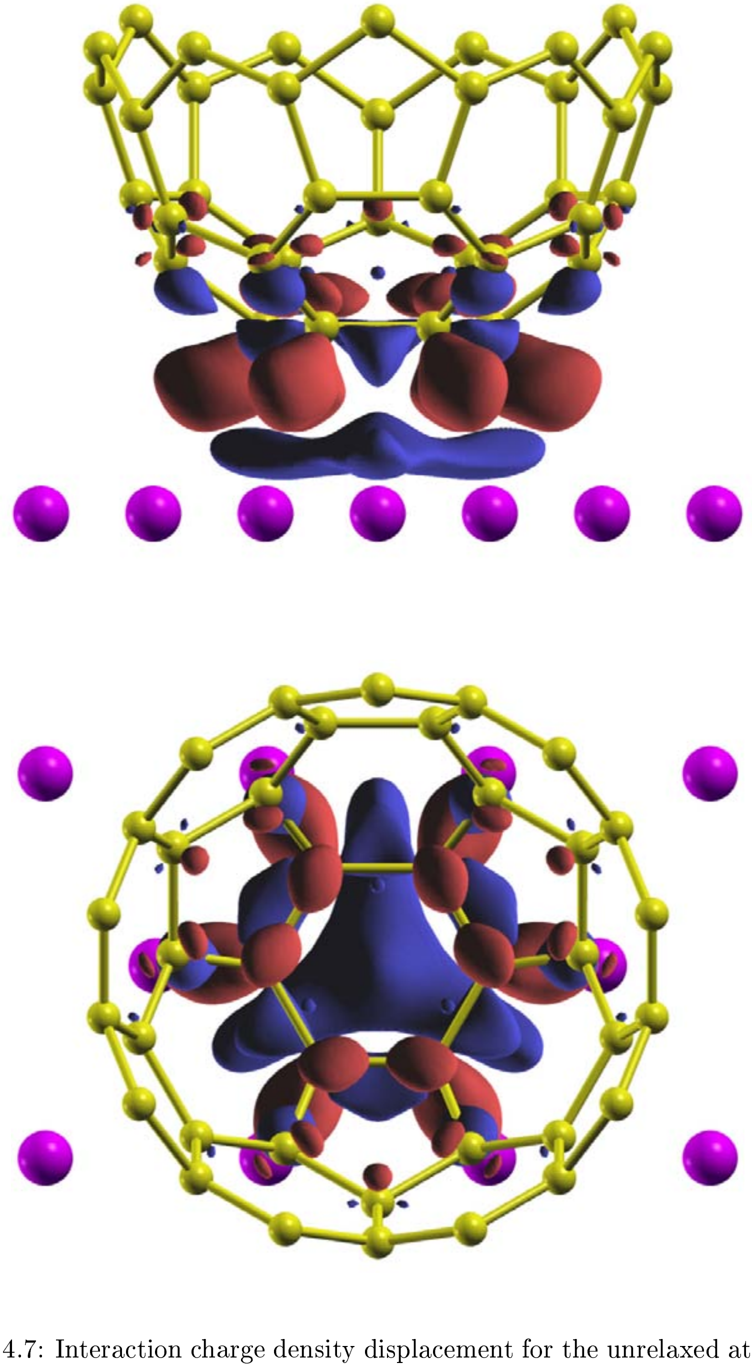


Figure 4.7: Interaction charge density displacement for the unrelaxed atop configuration, only the relevant part of the system is shown. Yellow balls and sticks represent C atoms and C-C bonds; magenta balls indicate Al atoms. Red and blue isosurfaces represent charge accumulation and depletion, respectively. Notice the partial sp^3 rehybridization of the lowermost C atoms.

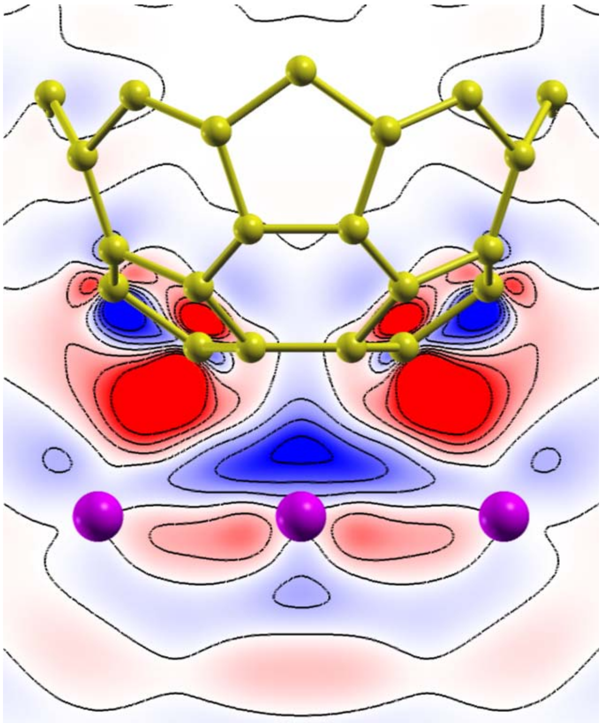


Figure 4.8: Same as Fig. 4.7, but charge density depletion/accumulation is represented as a contour plot for a vertical cutplane passing through the central surface Al atom.

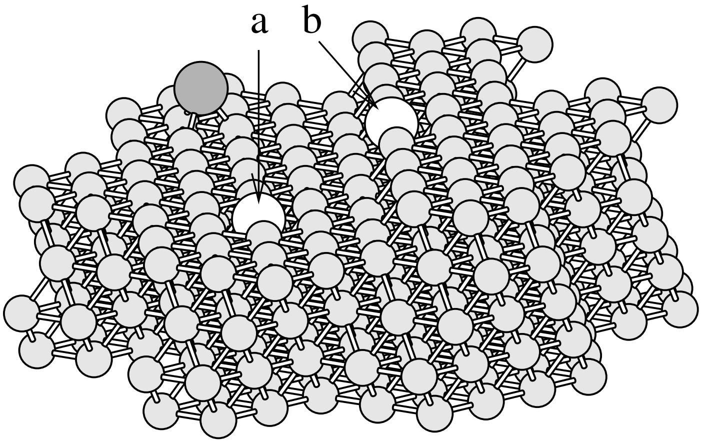


Figure 4.9: Adatom-vacancy pairs on a clean Al(111) surface. An isolated Al adatom (large dark grey ball) can be produced by leaving a vacancy (large empty ball) in a flat terrace (a) or at a monoatomic step (b).

reconstruction.

4.6 Adatoms and vacancies

We have already pointed out that the adatom-vacancy pair production on the clean Al(111) surface has a rather high energy cost. A well annealed Al(111) surface consists of flat (111) terraces separated by monoatomic steps, so there are two possible mechanisms for creating an adatom, as depicted in Fig. 4.9. The process where a vacancy is left in the monoatomic step has a lower energy cost (1.3 eV) compared to that of 1.7 eV relative to a flat terrace. However, both mechanisms are unlikely to occour at the reconstruction temperature of 490 K. We will show in the following that the presence of a C_{60} overlayer substantially reduces the energy cost for Al adatom production, due to the steric effect discussed in the previous section.

To see this, we start from a relaxed symmetric atop configuration, we remove the central Al atom and relax again the atomic structure of the molecule placed on the resulting surface vacancy (Fig. 4.10). This time the C_{60} binds to the substrate in a symmetric minimum configuration with six strong C-Al bonds whose length is 2.2 A. We notice that this bond length is identical to the optimized bond length values found in the hollow and bridge structure. At variance with those cases, here all C atoms belonging to the lowermost 6-ring can coordinate optimally to the surface Al atoms. As a consequence, the adsorption energy (2.34 eV) of C_{60} on the surface defect is much higher than all those obtained so far (Table 1).

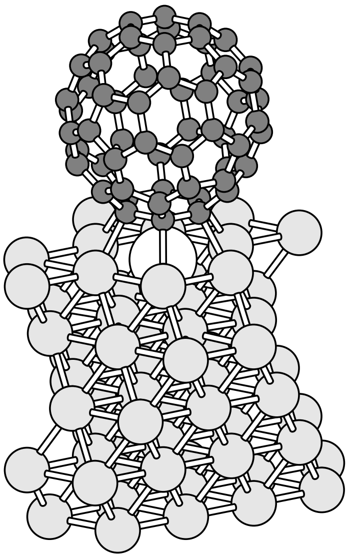
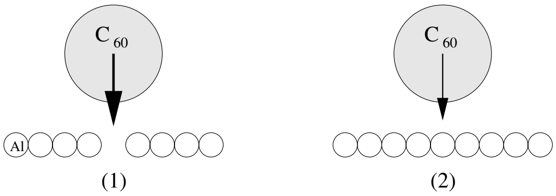
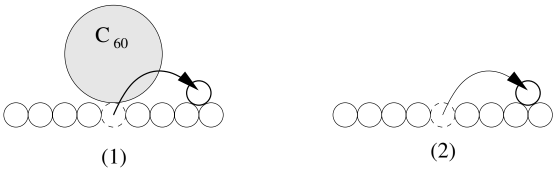


Figure 4.10: An Al vacancy (the missing atom is indicated as a large empty ball) can be created directly underneath an adsorbed C_{60} molecule, with a substantially lower energy cost.



Adsorption energy (E_{ads})



Vacancy/adatom pair creation energy (E_{ad-vac})

Figure 4.11: Comparison of adsorption energies (top) and adatom-vacancy pair production energy cost (bottom).

The meaning of this result is twofold, as we show schematically in Fig. 4.11. Considering the adsorption energies, C_{60} binds preferentially at vacancy sites. This could explain the experimental observation of occasional isolated islands which nucleate in the middle of flat terraces [51]. On the other hand, considering the adatom-vacancy pair production, this result confirms the significant lowering of the associated energy cost induced by the C_{60} overlayer, which is now 0.7 eV.

This value is still rather high, and one could wonder whether this mechanism applies at the experimental temperatures. We notice here that the calculated value of 0.7 eV corresponds to the total energy difference between an initial state where a C_{60} is adsorbed on a clean terrace, and a final state where the molecule sits on the vacancy while an isolated adatom is brought infinitely far onto a clean Al(111) terrace. In the real C_{60}/Al(111) system an Al adatom could interact with the adsorbates with even stronger covalent bonds (an undercoordinated adatom is expected in general to be chemically more reactive than surface atoms), and therefore contribute to a further lowering of the total energy.

To investigate this, we built a regular lattice of Al adatom-vacancy pairs by reintroducing the removed Al atom into the overlayer hollow site (in the middle of three adjacent fullerenes). After relaxation, the Al adatom ends up close to a bridging site between two C_{60} (corresponding to the positions evidenced in red for the (b) structure in Fig. 4.12). In this geometry the adatom can form one C-Al bond with each fullerene and two Al-Al bonds with the surface atoms. The computed total energy of this $(2\sqrt{3} \times 2\sqrt{3})R30^\circ$ structure is slightly lower (~ 0.03 eV) than that of the best unreconstructed geometry (*hollow-6*).

In summary, according to our calculations, the creation of Al adatom-vacancy pairs in the C_{60}/Al(111) system with $(2\sqrt{3} \times 2\sqrt{3})R30^\circ$ geometry has practically zero energy cost. In the following section we will investigate whether a partial relaxation of the translational symmetry (i.e. a larger unit cell) allows for a favorable energy balance.

4.7 (6×6) reconstruction

The results of the previous section indicate that the extraction of the Al atom occupying the central region below an adsorbed C_{60} is energetically *much* cheaper than the creation of a vacancy on the clean Al surface. Moreover, an adatom placed in the interstitial regions of the C_{60} overlayer interacts with the adsorbates forming strong C-Al covalent bonds. We thus investigate the possibility that a mechanism based on vacancy-adatom formation may be at the origin of the observed stable (6×6) reconstruction. We use a supercell containing 180 Al atoms (36 atoms per layer) and three C_{60} molecules for a total of 360 atoms and 1260 electrons which were represented on more than 800 wavefunctions. We create a vacancy under two of these, and let the corresponding pair of Al adatoms form a dimer, to take advantage of the associated binding energy which is 0.5 eV [55] on a clean surface. Then we study two different positions for the third C_{60} and the addimer, one with the Al dimer placed directly underneath the

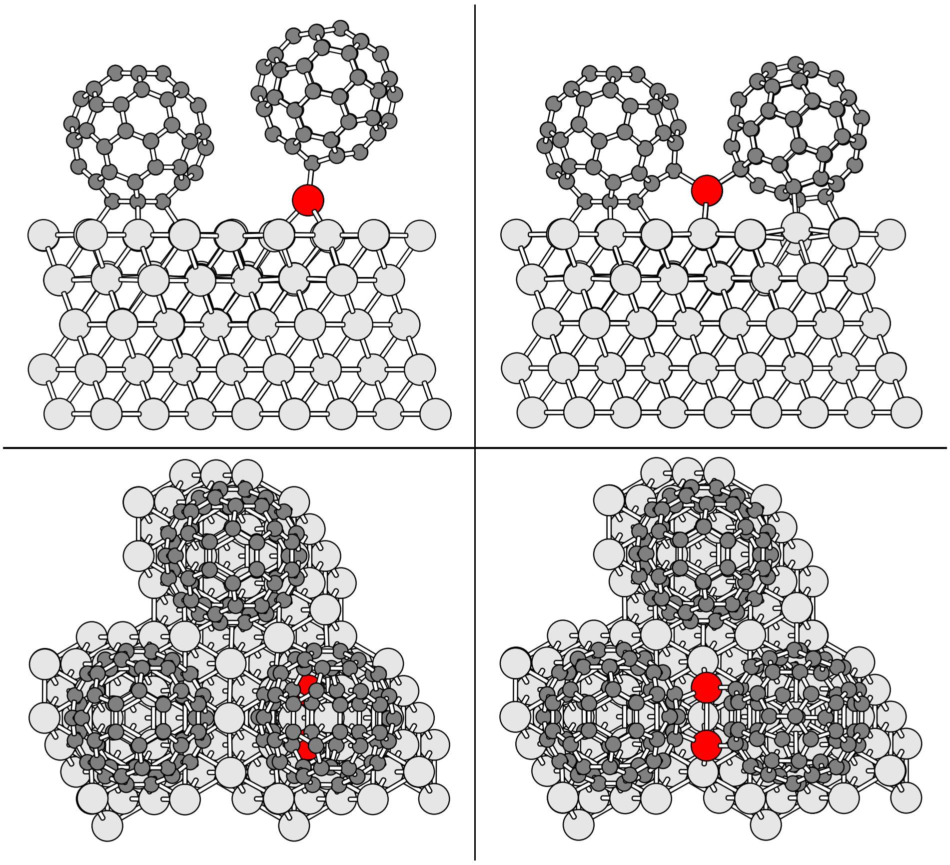


Figure 4.12: Relaxed atomic positions for the two (6 × 6) configurations considered in this work. Left: the Al addimer is placed directly underneath the "raised" C_{60}; right: the dimer is placed in the interstitial site of the overlayer, between a C_{60} sitting on a vacancy and the raised one which appears rotated (one single C atom facing the substrate). In the front views (upper panels of the figure) the C_{60} molecules lying in the background are not shown for clarity.

raised C_{60}, which is thus completely detached from the substrate, and a more "compact" one, with a less important vertical displacement but a substantially improved C_{60}-Al coordination. We fully relax both structures, which we will call (a) and (b) in the following, and obtain the equilibrium positions represented in Fig. 4.12. In both cases the positions of the two C_{60} molecules sitting on the vacancies (which we identify with the molecules appearing at a lower (L) height in STM topographs) have a structure which is quite similar to that already discussed in the previous section for the C_{60}/vacancy system, so we will concentrate our discussion on the equilibrium geometry of the "raised" molecule (R). The vertical height of R is different in the two cases (2 Å in (a), only 0.5 Åin (b)), so at first sight the structure which best fits the experimental findings is (a). However, in our calculations, configuration (b) is much more favorable energetically, providing a 0.8 eV gain with respect to the best unreconstructed structure, while (a) results in a positive energy contribution of 0.8 eV (i.e. the total energy of (a) is 1.6 eV higher than that of (b)). This energy difference can be ascribed to the fact that in (b) R is still in contact with the substrate, and therefore more C-Al covalent bonds can be formed, while in (a) R interacts only with the underlying Al dimer. In both cases the R cage is strongly distorted by the interaction with Al atoms, some C-C bonds appearing even more stretched than in the unreconstructed structures. This is in agreement with the otherwise paradoxal experimental observation that R molecules bind to the substrate more strongly than the L ones, despite the shorter distance of the latter from the Al surface.

As we remarked above, there is a discrepancy between our results and the experimental data, i.e. our most favorable (6×6) structure displays a protrusion of 0.5 Å, while a value of 1.9 Å was inferred from accurate STM measurements. Many factors could be responsible for this discrepancy, but we consider that the most likely ones stem from two sources. On one hand, the two-configuration sampling of the extremely wide configurational space for the (6×6) supercell is admittedly poor, and a more systematic search could reveal many other arrangements, possibly in agreement with our calculated energetics and the experimental geometry. However, because of the extremely high computational cost of such an investigation, and of the total absence of further experimental indications to countercheck our findings, we don't consider useful at the present stage to pursue this route. On the other hand, while the convergence of our results has been checked with respect to all tunable computational parameters, there remains an uncertainity due to our use of the local-density approximation, so it is important here to discuss how this approximation could affect our results.

4.8 LDA vs. GGA

LDA was found to give a reasonable description of C-Al bonding in previous studies. In order to verify whether GGA could be a better choice for our present work, we compared the performance of the two functionals in some selected

Configuration	LDA	GGA
atop	0.98	0.22
hollow-6	1.37	0.36
vacancy	2.34	0.73

Table 4.2: Adsorption energies (in eV) for selected equilibrium configurations in the $(2\sqrt{3} \times 2\sqrt{3})R30^\circ$ phase. The in-plane lattice constant was adjusted to the consistent Al theoretical bulk lattice parameter calculated within each approximation.

representative cases. For the adsorption on hollow-6, atop and vacancy sites GGA gives practically identical equilibrium geometries, with the same energy trend. However, the calculated GGA adsorption energies are very different from the LDA ones, as shown in Tab. 4.2. In particular, GGA predicts a C_{60}/Al(111) interaction which is much weaker than the estimate (cited in the introductory sections of this Chapter) for C_{60}/fullerite(111). This is in sharp disagreement with all experimental observations, and for that reason we used LDA for our study since it gives more realistic energies.

It is interesting to look how badly LDA and GGA perform in describing the C_{60}-C_{60} van der Waals interaction in solid fullerite. According to the calculations of [42] the LDA description is nearly optimal, with cohesion energy, bulk modulus and equilibrium distance in good agreement with experiment. We performed some tests and obtained substantially different conclusions, maybe thanks to a richer BZ sampling. A recent study [43] confirms our preliminary findings: the binding energy of C_{60} in bulk fullerite is around 1 eV within LDA, much lower than the experimental value of 1.73 eV. Local or semilocal exchange and correlation functionals are clearly not able to capture the physics of van der Waals bonding, and this fact is even more striking in GGA, which gives a vanishing minimum for the fullerite cohesion (< 0.1 eV), at a lattice parameter which is unphysically large.

It is not unreasonable to think that the wrong treatment of C_{60}-C_{60} interaction could be one of the reasons why the (a) structure is so much unfavorable with respect to (b). In fact, within the LDA approximation, in our calculations there is no compressional stress in the overlayer, so energy can be gained only by covalent interactions, and as a consequence the “wetting” of C_{60} by Al atoms (which is optimal in (b)) is strongly encouraged. The compressional stress in the overlayer could however play a further important role besides that on chemical bonding, e.g. it is the only reasonable justification for the observed long-range order in the (6×6) overlayer. Moreover, it could correlate strongly with temperature, i.e. a C_{60} monolayer at 500 K constrained to two-dimensional motion could have a substantially higher average lattice constant at thermal equilibrium than at RT. These factors play in favor of the (a) structure, but at present we are not able to quantitatively estimate their importance, nor to state whether they could eventually thermodinamically stabilize (a) with respect to (b).

4.9 Summary and outlook

To summarize the main results of this Chapter,

- the preferred site for the adsorption of the C_{60} molecule on the clean Al(111) surface is *hollow*, and not *ontop* as it was previously speculated;
- the energy of C_{60}/Al(111) is lowered significantly by an appropriate reconstruction of the substrate;
- strong Al-C covalent bonding is identified as the main driving force for the reconstruction;
- there is no need for the observed adatoms to migrate from steps, as they can be generated *locally* together with vacancies.

Possible future developments of the present work span many promising directions. Concerning C_{60}/Al(111), many features could be investigated to further refine the theoretical knowledge of this system. It would be interesting to evaluate some relevant kinetic barriers, e.g. for the C_{60} mobility on a clean surface, or for the production of an adatom-vacancy pair in presence of the overlayer. For example, it is not clear if the molecule translates with or without rotation during in-plane motion; the covalent character of the bonding would suggest a flat motion, perhaps a zig-zag hopping between neighboring bridge and hollow sites, which are energetically close and could therefore explain the observed high mobility. On the other hand, the presence of adsorbates is known to lower the barrier for vacancy production on clean Al(111), and it is quite reasonable to think that this could be true also for C_{60}. This would further strengthen the arguments in favor of our proposed adatom-vacancy reconstruction mechanism, but in order to obtain quantitative results, a full study taking into account all possible reaction paths would be necessary.

Perhaps the most stimulating development would be to transfer these ideas to C_{60} adsorption on other metallic substrates. Strong *lateral* interactions between C_{60} and substrate steps have been reported for example in C_{60}/Pd(110) [49]. Moreover, trapping of C_{60} in nanometer-sized vacancies has been reported on the Cu(111) surface [59]. Finally, the debate concerning the physical origin of the "bright-dim" STM contrast for C_{60}/Ag(100) is still unsolved, although recent findings tend to support overlayer-induced rearrangements of the Ag substrate.

The creation of irregular surface sites involving vacancy formation and strong lateral bonding between C_{60} and substrate (ad)atoms may thus be a quite general C_{60}/metal reconstruction feature. *Ab-initio* simulations will be an invaluable tool to investigate details of the C_{60}/metal interaction that are not directly accessible to the experiments, and thus give a significant contribution to the overall physical understanding of these mechanisms. Our hope is that the encouraging results we obtained in the present work will stimulate in the future similar studies about C_{60} adsorption on metals, a research field which is of wide interest and where much is still to be done from the theoretical point of view.

Chapter 5

Magnesium Oxide on Ag(001)

5.1 Introduction

Ionic insulators like simple metal oxides are routinely used as inert supports for catalytic metal clusters. Unfortunately, the use of an insulating substrate severely hampers the experimental characterization of such systems by means of standard STM and spectroscopic techniques, due to sample charging problems. In order to circumvent this difficulty, it has become common practice to replace the insulating substrate by a thin oxide film deposited on a metallic surface. In this way, the advantages provided by the ceramic material are not lost (a thickness of ~ 10 atomic layers is generally sufficient to reproduce the chemical and electronic properties of the oxide surface), while the electrons can now tunnel through the film, thus eliminating the unwanted build up of charge at the sample surface.

The interest in thin oxide films is steadily growing also for other reasons, related to (nano) electronics. On one hand, they are necessary as dielectric materials, e.g. in MOS-FET transistors, whose size is progressively decreasing as a consequence of technological miniaturization trends. On the other hand, they serve as tunnelling barriers in several nanodevices and therefore they have to be extremely thin to guarantee sufficiently large currents at low applied voltages.

Finally, there is considerable fundamental interest in understanding the physics of ultrathin oxide films. The reduced dimensionality, together with the interaction with the substrate, lead to chemical and physical properties that are often very different from those of the bulk crystal surfaces. In the experimental characterization of such systems, STM and STS play often an important role. However, in many cases of interest, the electronic and structural features are not easy to disentangle. In order to rationalize these results and help unvealing the underlying physical mechanisms, *ab-initio* simulations can be an extremely valuable tool, as we will demonstrate for the case of the MgO/Ag(001) system.

This chapter is organized as follows: in Section 5.3 we will study the structural properties of unsupported MgO(100) thin films, while their electronic properties will be discussed in Section 5.4. This information will be valuable in our study in order to separate the effects coming from the MgO film thickness from those arising from the interaction with the Ag(100) substrate. In Section 5.5 we will briefly introduce the electronic properties of the clean Ag(100) surface, and finally in Section 5.6 we will discuss our results for the MgO/Ag(100) system, starting with an overview of the existing literature on the subject.

5.2 Computational details

The present study has been entirely done with the ABINIT code [1]. For the study of 0-3 ML supported MgO films, we modeled the Ag surface by 9 layer slabs in a periodic supercell geometry, separated by vacuum layers at least 15 Å thick, and with an in-plane lattice parameter corresponding to the calculated equilibrium value for bulk Ag. The MgO(100) films were adsorbed on both sides of the slab, in order to have symmetric systems and thus avoid electric fields. The atomic positions of Mg and O atoms, and those of the Ag atoms on the two outermost Ag layers, were fully relaxed (to a tolerance of 5×10^{-4} Ha/Bohr) by auxiliary calculations with a 5 layer Ag slab, while the other positions were set to ideal bulk Ag coordinates. Interfacial O atoms are placed on top of the Ag(100) surface atoms, which is the most favorable configuration according to our calculations, in agreement with the literature. We use an 8×8 Monkhorst-Pack k-point grid to sample the surface Brillouin zone (SBZ) [2], corresponding to 10 special points in the irreducible wedge. This k-point set was augmented to 16×16 for the calculation of the surface LDOS. The calculated total energies and lattice constants are well converged for a plane wave cut-off of 70 Ry. All pseudopotentials were generated with the FHI98PP code [56], and thoroughly tested for transferability. Non-linear core corrections were added for the Mg atom in order to obtain reliable structural properties for magnesium oxide.

Throughout this work we used the LDA approximation for the exchange and correlation functional. We checked also the GGA performance in this study, and while we could obtain slightly improved structural properties, we found less accurate results for the electronic structure (which is the main goal of our study).

5.3 MgO thin films: structural properties

We present here a detailed study of the structural properties of thin MgO films with 1-5 monolayer thickness.

---

[1] The ABINIT code [113] is a common project of the Université Catholique de Louvain, Corning Incorporated, and other contributors (URL http://www.abinit.org).

[2] For pure MgO (insulating) systems a 4×4 sampling was largely sufficient to ensure convergence.

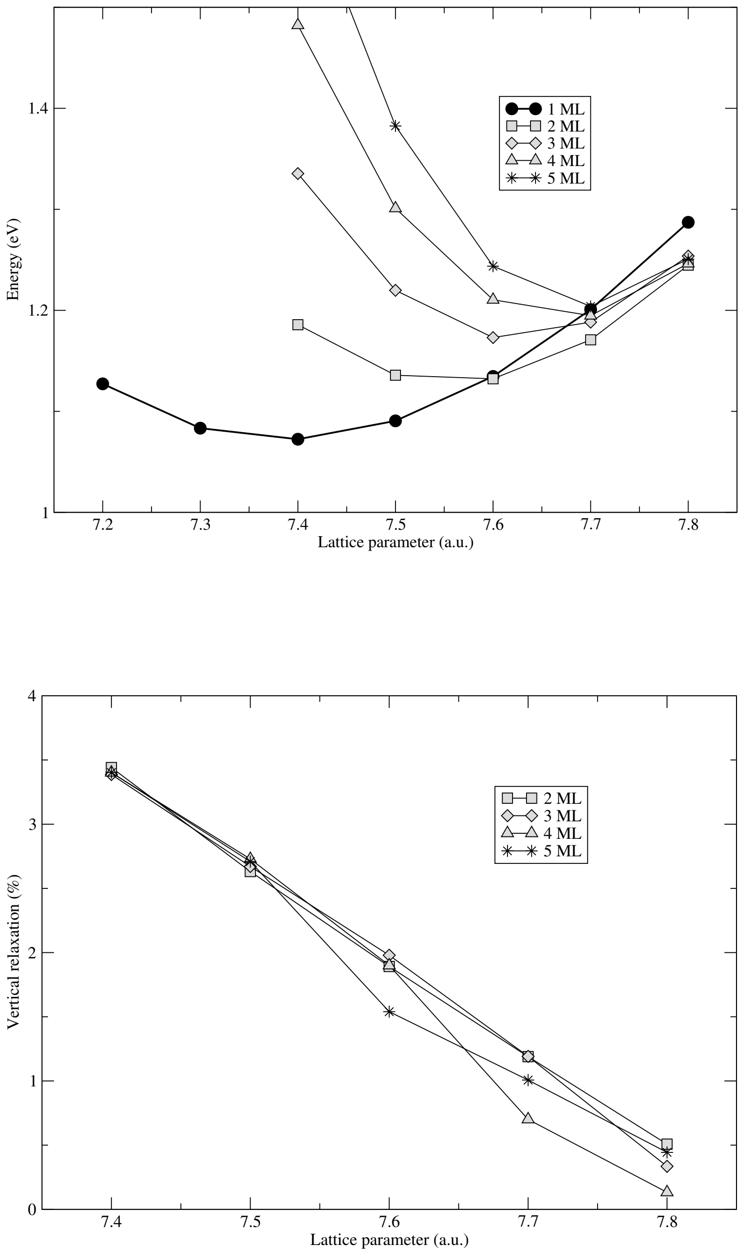


Figure 5.1: LDA energy vs. in-plane lattice parameter for unsupported MgO films (upper panel). In the lower panel the vertical relaxation of the slab is given as a function of the in-plane lattice constant (relaxation in percent relative to the bulk equilibrium lattice parameter).

Thickness (ML)	a_0 (a.u.)	Contraction (%)
1	7.39	5.6
2	7.56	3.3
3	7.63	2.4
4	7.67	1.9
5	7.69	1.6

Table 5.1: In-plane equilibrium lattice constant of isolated MgO films.

While restricting our analysis to the standard rocksalt structure, we varied the in-plane lattice constant and we fully relaxed the atomic positions in each case. We report in Fig. 5.1 the results for the energy as a function of the in-plane lattice parameter. In the same figure we show the relaxation along the perpendicular axis as a function of the in-plane strain. The equilibrium in-plane lattice constants and the percentual reduction with respect to the bulk value ($a_0 = 7.82$ a.u.) are summarized in Table 5.1. Thicker slabs recover progressively the equilibrium lattice constant of the bulk, but the convergence with respect to the number of layers is rather slow, so that even a 5 layer film is contracted by ~ 1.5 % with respect to the bulk. These results can be related to some extent to the experimental observations for the MgO/Ag(100) system. The authors of Ref. [65] find that the relaxation of the compressional stress in the film by misfit dislocations between substrate and adsorbate starts to be important at a critical thickness of 3-6 ML. Here we show that the in-plane equilibrium parameter of 1-5 MgO films is contracted with respect to the bulk lattice constant, and this should lead to a good matching with the Ag substrate for a film thickness ≤ 4 layers, which is in good agreement for the experimentally estimated critical thickness. A qualitatively similar behaviour has been observed also for MgO deposited on other metallic substrates, e.g. Fe(001) [82], and could be related to the same physical property.

The in-plane contraction of ultrathin MgO films can be explained by simple arguments based on the predominantly ionic character of the MgO cohesion. In the simplest Born-Mayer model of an ionic crystal, the equilibrium is determined by the interplay of Coulomb interactions and short-range repulsion of the ion cores. If we assume rigid ion cores (which is a very simplistic approach to the description of MgO, but will suffice for the present discussion), and a purely exponential nearest-neighbour repulsive interaction, we can write the energy per atom as [87]:

$$U = z\lambda exp(-R/\rho) - \frac{\alpha q^2}{R}, \tag{5.1}$$

where z is the numbers of neighbors and α is the Madelung constant. The equilibrium lattice parameter depends in this model only on the ratio z/α, and in particular it is an increasing function of this ratio. For a squared monolayer, as for an hexagonal one, z/α is smaller than in the bulk rocksalt structure, and this roughly justifies the trend for the 2D structures to be tighter than the 3D bulk.

At a fixed lattice constant of 7.82 a.u. the formation energy of 2-5 MgO films is identical within 0.01 eV. This means that the MgO(100) surface energy is converged within 0.005 eV already for a 2 ML slab, as already reported by [76]. In our LDA calculations we find a value of 0.62 eV per surface MgO unit. Within this fixed value for the in-plane lattice constant the effects of ionic relaxation are very small, and amount to 0.02 eV for the 5 layer slab. The surface "rumpling" (slight height difference between the surface Mg and O ions) is also rather small, and in all systems it amounts to 0.04-0.05 A, which is in agreement with previous theoretical [76] and recent experimental [83] data.

5.4 MgO electronic properties

The underestimate of the electronic gap in non-metals is a known problem of LDA, and is due to the fact that the one-particle eigenvalues enter the formalism as Lagrange multipliers and in principle have no direct physical meaning. Experience indicates that the discrepancy with respect to the experimental spectra is in many cases limited to a rigid shift of valence and conduction bands (the so-called "scissor-operator"), and the qualitative trends are relatively well reproduced. Previous theoretical results for the bulk quasiparticle spectra obtained within the GW approximation suggest that this is the case for bulk MgO [75, 77].

The electronic structure of the MgO(100) surface has been studied by many authors, and the physical nature of the unoccupied surface bands remained somewhat controversial for a while. Very recently, the system was studied by means of the GW method, and important insight could be gained from the theoretical point of view [115]. These results indicate that for the MgO(100) surface the "scissor operator" model must be used with some care.

5.4.1 Bulk

We report in Fig. 5.2 the LDA band structure calculated for bulk MgO along the high simmetry lines in the Brillouin zone (BZ). The calculated electronic gap is 5.2 eV, i.e. strongly underestimated with respect to the experimental value (7.7 eV), and in line with the literature values [77, 75]. It is interesting to discuss briefly the orbital character of the energy bands depicted in Fig. 5.2. The valence bands can be associated unambiguously to O $2s$ and O $2p$ orbitals, for the structure at ~ -17 eV and ~ 0 eV respectively. On the other hand, there has been some debate concerning the MgO conduction band, which in the naive charge-transfer model of ionic materials is commonly associated to cation empty orbitals. In Ref. [78] this point is treated extensively, and the authors find this picture inadequate for MgO, since the conduction-band bottom has a free-electron character with considerable oxygen $3s$ components. Our calculations support their findings. MgO conduction-band states at the Γ point have a nearly

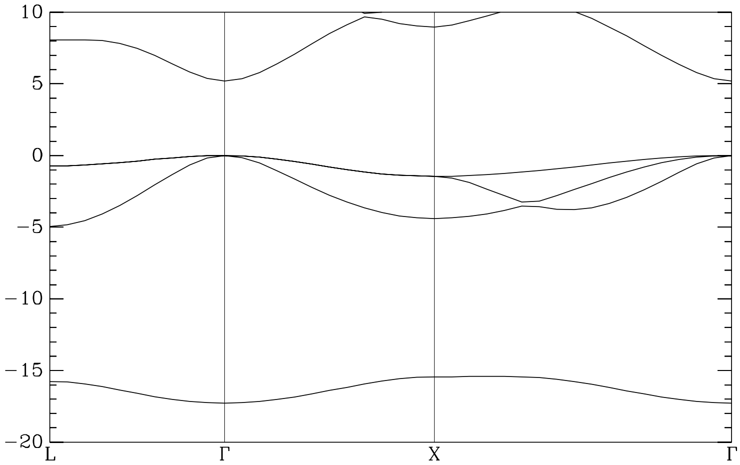


Figure 5.2: LDA band structure for bulk MgO. Energies are in eV, the zero is set to the valence band top.

free-electron-like wavefunction with spikes on the ions due to the orthogonality constraints to the valence states.

5.4.2 MgO(100) surface

We show in Fig. 5.3 the LDA band structure for a 7 ML MgO(100) slab, which represents well the semiinfinite surface, compared with the projection of the bulk band structure along the (100) direction. The (100) surface shows a reduction of the fundamental band gap with respect to the bulk value. This result is already known both theoretically [75] and experimentally [84], and is due to the presence of an unoccupied surface band in the bulk energy-gap region. However, the details of the fundamental physical mechanisms governing the nature of this surface state are still not clear [79]. This is of great importance for correctly interpreting the experimental results for the MgO/Ag(100) system, so we will discuss this point in full detail.

A possible explanation for the surface band gap reduction is reported in a recent review [74], and is based on electrostatics. The undercoordinated atoms at the surface experience a Madelung potential which is smaller than that in the bulk. This is interpreted as a perturbation on the effective potential felt by the electronic orbitals located at the surface sites. The author shows that this leads to a lowering of the eigenvalues associated to the conduction band, and this would explain the observed surface band gap reduction. We find this interpretation quite unsatisfactory, for two main reasons. First of all, the reduction in the Madelung constant (-4 % with respect to the bulk) at the close-packed (100) surface of rocksalt ionic materials is relatively small [3], and thus we expect

---

[3] The value of 4 % is calculated for an ideal termination of the bulk, and atomic relaxation

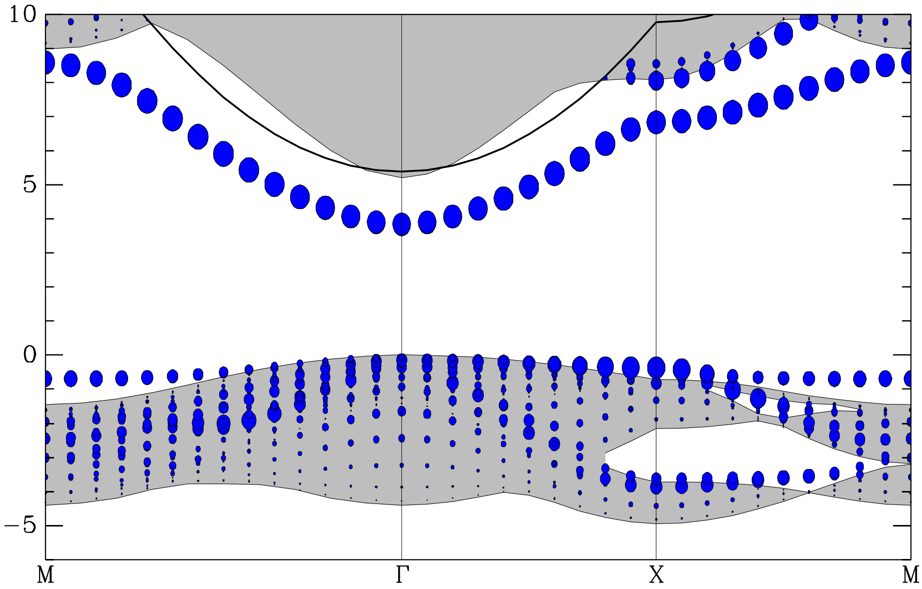


Figure 5.3: Calculated band structure for the MgO(100) surface. Shaded areas represent the 2D projection of the bulk MgO bands along the (100) direction, aligned with respect to the valence band top; thick continuous line indicates the energy dispersion of a free electron in vacuum, i.e. the bottom of the parabolic dispersion is the calculated vacuum level. The radius of the filled circles is proportional to the probability of the state to be outside the outermost atomic layer. Energy values are in eV.

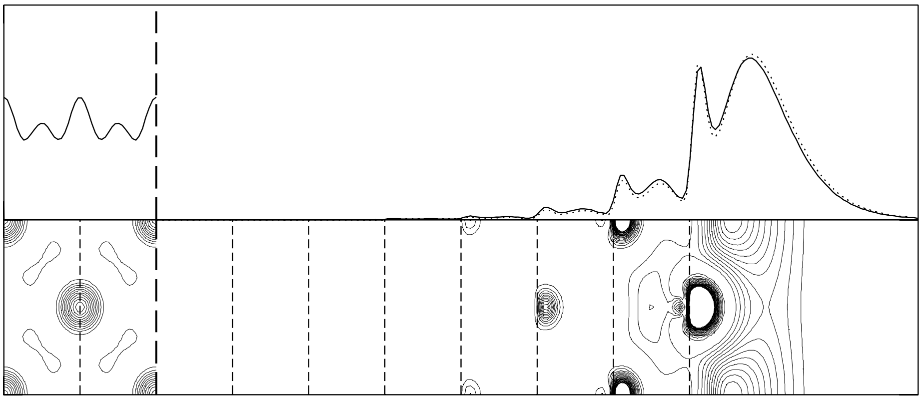


Figure 5.4: Surface conduction state at Γ for MgO(100). Both contour plots (lower panel) and xy plane averages (upper panels) refer to the charge density of the state. The left part of the figure shows the lowest conduction state in the bulk. Dashed lines indicate the positions of the atomic layers. We used a symmetrical 15 layer slab, with relaxed atomic positions. The dotted line shows the lowest conduction state for the unrelaxed surface.

it to produce a rather weak perturbation of the eigenenergies of the bulk-derived states. Furthermore, this argument should apply even better to the oxygen $2p$ orbitals (which are much more localized than the diffuse Mg $3s$), while we can see in Fig. 5.3 that there is no occupied surface state above the bulk valence band edge.

Other authors proposed that the gap reduction was due to the Stark-shift induced on the unoccupied cation orbitals (Mg $3s$) by the electric fields at the surface. Again, this interpretation makes sense only in a tight-binding picture for the MgO conduction band, which is clearly not appropriate.

A few months ago, M. Rohlfing *et al.* [115] provided theoretical evidence that the unoccupied surface states observed in MgO(100) and LiF(100) are the $n = 1$ terms of the infinite serie of image-potential-induced states (IPS), which are predicted to exist at the surface of *any* material. They show that the wavefunctions of these surface states are qualitatively similar to the corresponding LDA empty states (i.e. most of their weight is outside the material, see Fig. 5.4), but their behaviour is strongly influenced by long-range dielectric screening, an effect which is absent in standard DFT approaches.

In the following sections we will introduce the physics of IPSs, and then we will analyze the role played by these ideas in the study of the MgO/Ag(100) system.

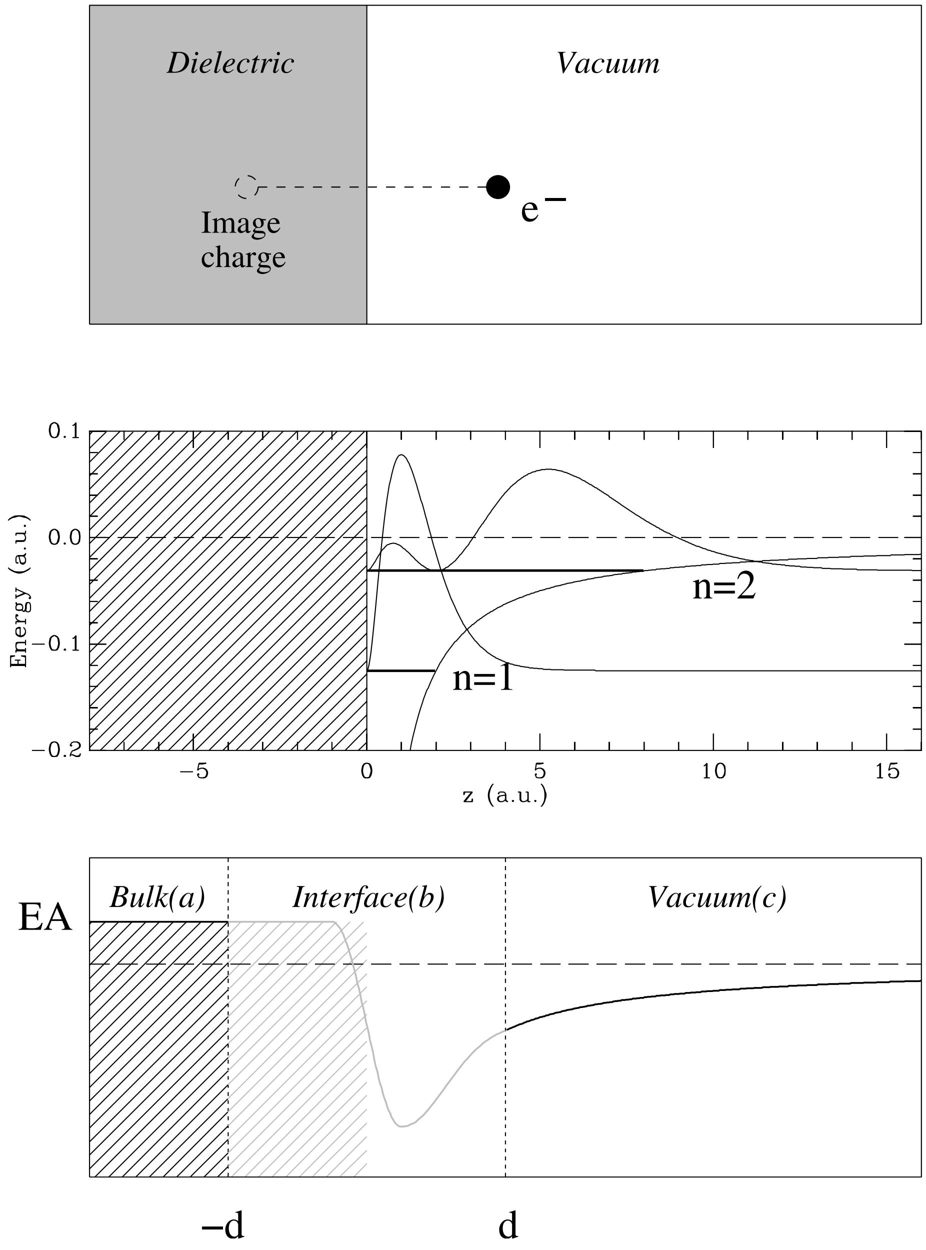


Figure 5.5: Physical models for the IPS. Upper panel: the dielectric response of the medium generates an attractive potential for an electron; central panel: simple 1D quantum mechanical description in terms of hydrogen-like wavefunctions; lower panel: for a realistic material the bulk electronic affinity EA (a), and the atomistic details at the interface (b) must be taken into account.

5.4.3 Image-potential-induced surface states

Given a semiinfinite homogeneous medium with a flat surface and dielectric constant ϵ, an electron in the vacuum at a distance z from the surface experiences an electrostatic potential given in atomic units by [88]

$$V_{im}(z) = -\frac{1}{4z}\frac{\epsilon - 1}{\epsilon + 1}. \tag{5.2}$$

By representing the medium as an infinite potential barrier we can exactly solve the one-dimensional Schrödinger equation for the electron [93]

$$\Big(-\frac{1}{2}\frac{d^2}{dz^2} - V_{im}(z)\Big)\phi_n(z) = E_n\phi_n(z), \qquad \phi_n(0) = 0 \tag{5.3}$$

which is identical in form to the radial Schrödinger equation for $r\psi$ for a Coulomb potential due to a nucleus of charge Z with $l = 0$, where

$$Z = \frac{1}{2}\frac{\epsilon - 1}{\epsilon + 1} \tag{5.4}$$

and has hydrogenic solutions with eigenenergies $E_n = -Z^2/2n^2$. Assuming a flat surface, the system is invariant with respect to translations in the xy plane, thus these bands have a free-electron dispersion parallel to the surface, and the full solution can be written as

$$\psi_{\mathbf{k}_{||},n}(\mathbf{r}) = e^{i\mathbf{k}_{||}\mathbf{r}_{||}}\psi_n(z), \qquad E_{\mathbf{k}_{||},n} = \frac{k_{||}^2}{2} + E_n \tag{5.5}$$

While this model accounts for the correct asymptotic limit of the potential in the vacuum, it is inadequate for $|z|$ less than some value d, which is of the order of one interatomic distance. Moreover, the infinite potential barrier is clearly an oversimplification for the electron scattering properties of a given material.

The description of the $z < -d$ (a) region can be qualitatively improved by the knowledge of the bulk band structure. Generally speaking, if the bulk conduction band (CB) lies *above* the vacuum level (negative EA), IPSs decay exponentially in the material as $\sim e^{\frac{z}{\sqrt{E_{CB}-E_n}}}$, while if the CB edge lies *below* vacuum (positive EA) some (or all) IPSs appear as resonances in the continuum of the bulk-derived Bloch waves. Most importantly, in the case of *negative* EA, the presence of IPSs determines a significant reduction of the fundamental electronic band gap, and this would explain the important surface gap reduction discussed in the previous section for MgO(100) [84].

For an accurate estimate of the effective potential in the interface (b) region an accurate first-principles description at the atomistic level is required in each case. This is quite important since the $n = 1$ state, which is most easily detected experimentally, and which determines the *surface fundamental gap* for many negative electron affinity insulators (like MgO), is in general strongly influenced

---

is expected to further reduce this value [74].

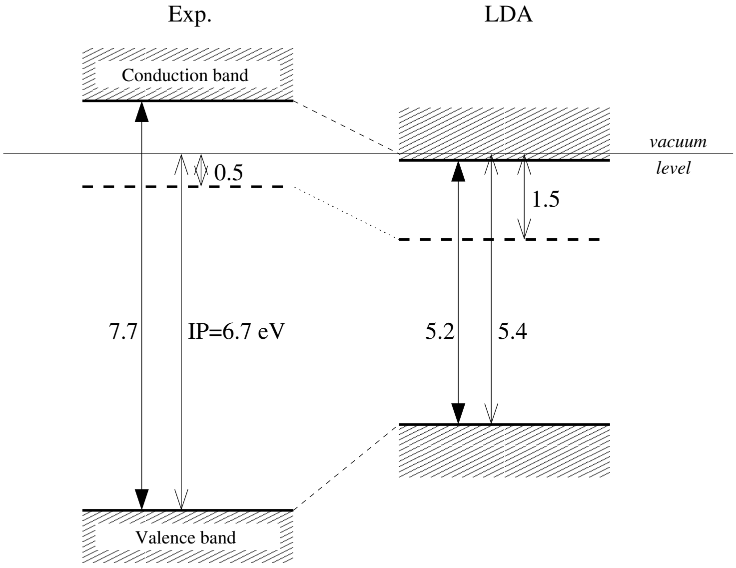


Figure 5.6: Electronic structure of bulk MgO. According to the available data MgO(100) has negative electronic affinity, i.e. bulk conduction band lies above vacuum level. LDA calculations (this work) underestimate the fundamental gap and ionization potential, and give practically vanishing EA. Experimental data are taken from [81, 85, 84].

by the short-range details of the effective potential, and its properties are often difficult to reproduce by means of simple macroscopic models [4].

5.4.4 IPSs on MgO(100)

In Fig. 5.6 we summarize the results for the MgO bulk and surface electronic structure (at the Γ point) obtained in this work within LDA, compared with the experimental picture. At a first glance, we can notice that the *difference* between the bulk and surface gap is almost identical in both cases. This fact could be interpreted as a proof of the reliability of LDA, in the context of the widespread "scissor-operator" model. We think that this interpretation is somewhat misleading, and that the good agreement of the LDA surface gap shift with experiment is mostly accidental.

Since we interpret the MgO(100) conduction band edge as an IPS, the parameter that is most relevant in our discussion is the *binding energy* of this state with respect to the vacuum level. This value is a direct signature of the

---

[4] On many metallic surfaces, accurate energies and wavefunctions for the image states could be obtained by a simple one dimensional model potential with only four adjustable parameters [94, 95]. For a ionic insulator the physical properties are substantially different and this model cannot be transferred straightforwardly.

deepness of the effective potential well (see Fig. 5.5) at the interface, which is of paramount importance in determining how far the center of gravity of the wavefunction lies away from the surface [5].

In this context, the IPS appears strongly overbound by LDA, 1.5 eV below vacuum level compared to the experimentally estimated value of 0.5 eV. Interestingly, this can not be due to the wrong LDA long-range potential in the vacuum region, since the substitution of the LDA exponential tail with an $\sim 1/r$ one would lead to a further lowering of the energy eigenvalue associated to this state. A possible explanation could be formulated in terms of nearly-free electron theory. LDA underestimates the position of the bulk conduction band edge with respect to the vacuum, which represents the effective potential barrier felt by an additional electron at the interface (see Fig. 5.5). This leads to an increase of the penetration length of the wavefunction into the material, and gives a negative contribution to the eigenenergy of this state. However, we expect this effect to be small, since already at the LDA level the wavefunction is centered in the outer portion of the interface (see Fig. 5.4), and a higher potential barrier in the bulk would lead to a small perturbation only.

We can thus speculate that the main source of error could be related to a wrong treatment of the effective potential felt by excited states at the interface. A further investigation of this problem would need to go beyond LDA, and here we will limit ourselves to a qualitative analysis of the deviations introduced by this approximation with respect to what one should expect for the true quasiparticle wavefunction. The main LDA errors in the effective potential can be summarized as follows:

- *lower* potential barrier in the bulk (a);
- *deeper* potential well at the interface (b);
- *exponential* instead of $1/r$ decay in vacuum (c).

All these effects operate in the same direction, i.e. lead to a spurious shift of the wavefunction importance towards the material.

This observation will be important when comparing our results for the MgO/Ag(100) system to experimental mesurements. We start in the next section with a discussion of the electronic structure of unsupported MgO films.

5.4.5 MgO(100) unsupported thin films

We summarize in Table 5.2 the results for the LDA electronic structure of MgO(100) unrelaxed (ideal bulk positions) and relaxed thin films (we relaxed the z coordinate while keeping fixed to the bulk value the in-plane lattice constant) . As a general trend, the gap is progressively reduced with decreasing film thickness, and the rumpling of the surface layer increases its value by ~ 0.3 eV. In order to separate the behaviour of the valence and conduction states we show

---

[5] See for example [101] for an extensive discussion on the relation between the spatial extent of image states and their binding energy

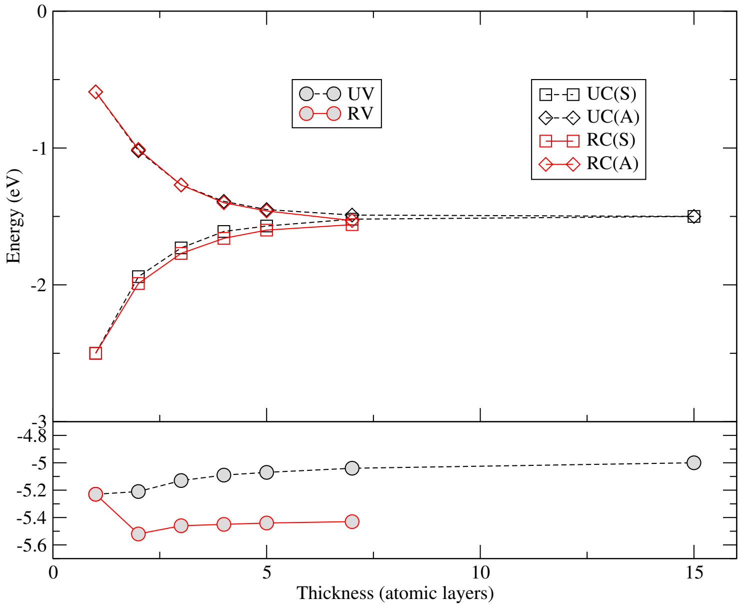


Figure 5.7: Summary of the evolution of the electronic structure as a function of film thickness. Upper panel: symmetric (S) and antisymmetric (A) conduction states at Γ; lower panel: valence band edge. Black dashed lines: atoms fixed in ideal bulk positions (Unrelaxed); red solid lines: Relaxed structures.

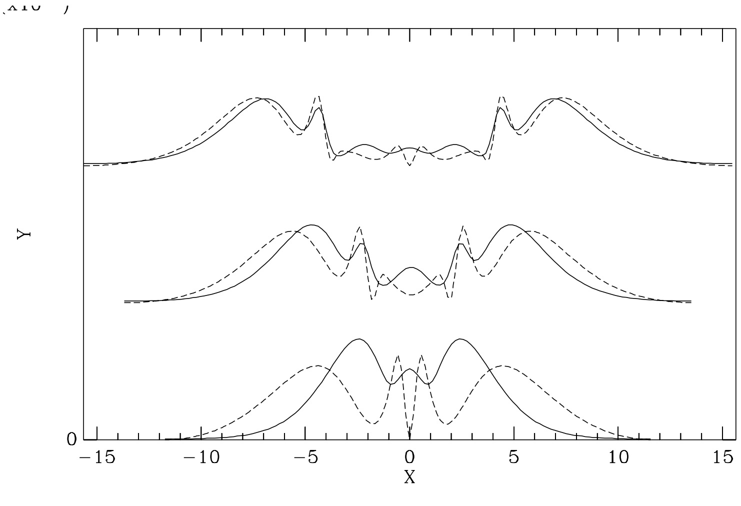


Figure 5.8: Charge density for the excited states of 1-3 ML thick unsupported films. Solid lines: symmetric, dashed: antisymmetric state. Distances are reported in a.u.

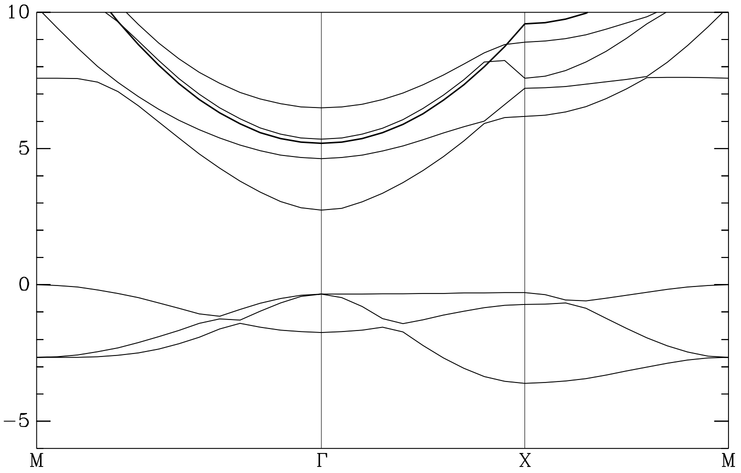


Figure 5.9: Band structure for the 1 ML unsupported MgO film. The energy dispersion of the lowest free-electron-like vacuum states (see also Fig. 5.3) is indicated as a thick line; the valence band edge is aligned to zero.

Thickness (ML)	Gap (eV)	
	Unrelaxed	Relaxed
1	2.73	-
2	3.27	3.53
3	3.40	3.69
4	3.46	3.79
5	3.49	3.84
7	3.50	3.87
15	3.50	3.89
Bulk	5.20	-

Table 5.2: Calculated electronic gap as a function of film thickness and relaxation.

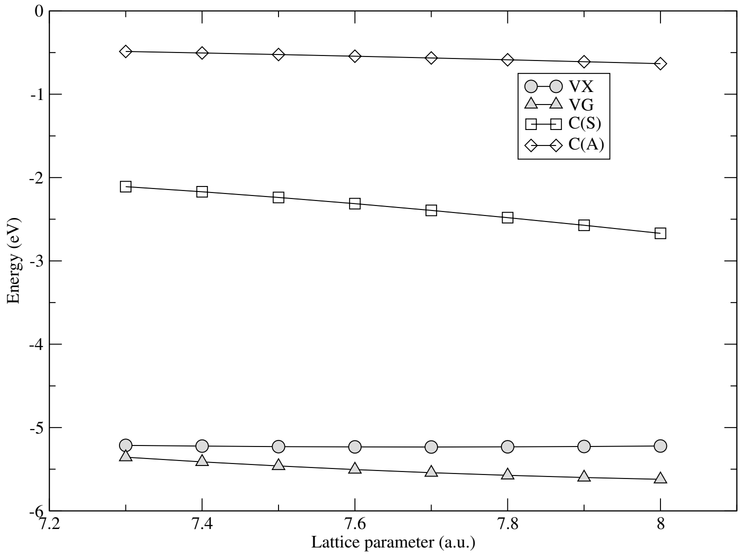


Figure 5.10: Evolution of the electronic structure of 1 ML film as a function of the in-plane lattice constant. We report the eigenvalue for the valence band top at the M point (VX) and at the Γ point (VG), and the symmetric (S) and antisymmetric (A) excited states at Γ.

in Fig. 5.7 the absolute energies of the corresponding eigenvalues with respect to the vacuum level. The valence band edge of the unrelaxed films shows little variation with thickness, while the surface rumpling leads to a significant increase of the ionization potential. On the other hand, the two lowest conduction states at Γ are marginally affected by surface relaxation. They are degenerate in the limit of high thickness, while the difference between their eigenenergies increases as the thickness is reduced. This is in agreement with our interpretation of the surface state as an electron bound to a 1D quantum well in the direction perpendicular to the surface. In a thick MgO film the two potential wells corresponding to opposite surfaces are well spaced apart, and the corresponding bound states do not interact. In thinner films the exponential tails overlap, so that there is a nonzero "hopping term" which splits the degeneracy in a symmetric and antisymmetric state (see Fig. 5.8). As a consequence, the symmetric state lowers its energy, thus reducing the electronic gap. From the discussion in the preceding subsection we expect this splitting, that is particularly important in the case of the monolayer, to be strongly exaggerated by the LDA overbinding of the surface state. In Fig. 5.9 we report the complete band structure for the 1 ML film. We notice that the band gap in the 1 ML case is not direct, because the valence band top lies at the M point of the SBZ.

Since we will deal with MgO films whose in-plane lattice parameter is not *a priori* fixed, because the interaction with the substrate could have a subtle interplay with the above mentioned structural results, it is interesting to investigate whether the electronic gap of the MgO films is sensitive to the in-plane strain (as much as the bulk electronic gap in MgO is sensitive to pressure). We show in Fig. 5.10 the energy of the relevant states which determine the electronic gap of the 1 ML film for increasing values of the in-plane lattice constant. The ionization potential is remarkably unaffected by compressional stress, while the main contribution to the gap variation is given by the shift of the conduction band towards higher energies. This leads to an increase of the surface gap by 0.14 eV when going from the calculated bulk lattice parameter of MgO (7.82 a.u.) to that of Ag (7.65 a.u.).

5.4.6 Unsupported films: a dielectric continuum model

In Section 5.4.4 we have shown the difficulties one encounters when trying to describe excited states of an insulating system on the basis of Kohn-Sham LDA eigenvalues of unoccupied orbitals. In order to check the reliabity of the qualitative trend given by LDA (gap narrowing for thinner films) we find it useful to study a simplified one-dimensional model of the film which takes into account the physical aspects of the problem that are neglected in LDA, i.e. the long-range dielectric response.

An electron outside a semiinfinite dielectric feels the classical image potential $V(z) = -\beta/4z$. If we consider instead a film of finite thickness t, the Laplace equation for the electrostatic potential in vacuum can be solved with the Bessel functions method [88] and has the solution:

$$V(z) = -\frac{\beta}{4z} + \frac{1-\beta^2}{4\beta} \sum_{n=1}^{\infty} \frac{\beta^{2n}}{z+nt} \tag{5.6}$$

Inside the film we assume that the potential equals the opposite of the the bulk electronic affinity (EA) of the dielectric, and that it does not depend on thickness. In order to avoid unphysical divergences in the interface region we choose a cutoff value r_{cut}, and between $z = 0$ and $z = r_{cut}$ we interpolate with a smooth curve as shown in Fig. 5.11 [6]. Then we solve the one-dimensional Schrödinger equation for an electron in the potential $V(z)$, and we obtain the binding energy for the first image state as shown in the inset of Fig. 5.11. This model gives a monotonic decrease in the binding energy of the $n = 1$ image state of the film for decreasing film thickness, contrary to the LDA results.

The dielectric continuum model is far too simple to be useful in practice, since the atomistic details of the system are completely neglected here, but it tells us that an important physical ingredient - the dielectric response of the film - provides a contribution which is opposite in trend with respect to DFT-LDA calculations. The message is that the analysis of LDA unoccupied states should be done with great care, and the results critically evaluated.

---

[6] r_{cut} and the shape of the interpolating curve are kept constant while thickness is varied.

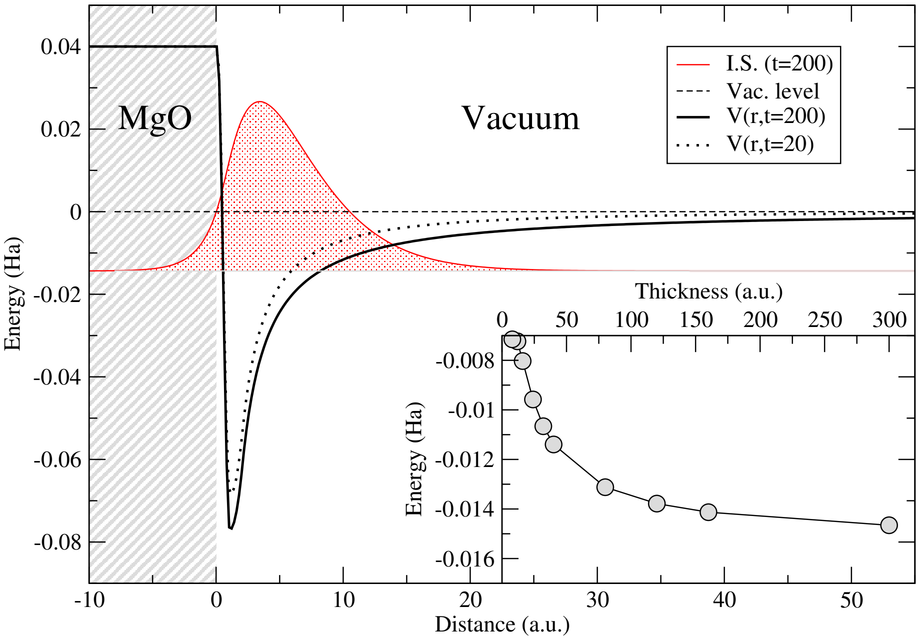


Figure 5.11: Dielectric continuum model for a thin MgO(100) film. The effective potential (Eq. 5.6) is shown for a 200 a.u. film (thick solid line) and a 20 a.u. film (thick dotted line). Vacuum level is indicated as a dashed line; inside the dielectric (left) the potential is set to -EA=1 eV. The first bound state for the 200 a.u. slab is shown in red. The insert shows the evolution of the eigenenergy of the lowest image state as a function of thickness.

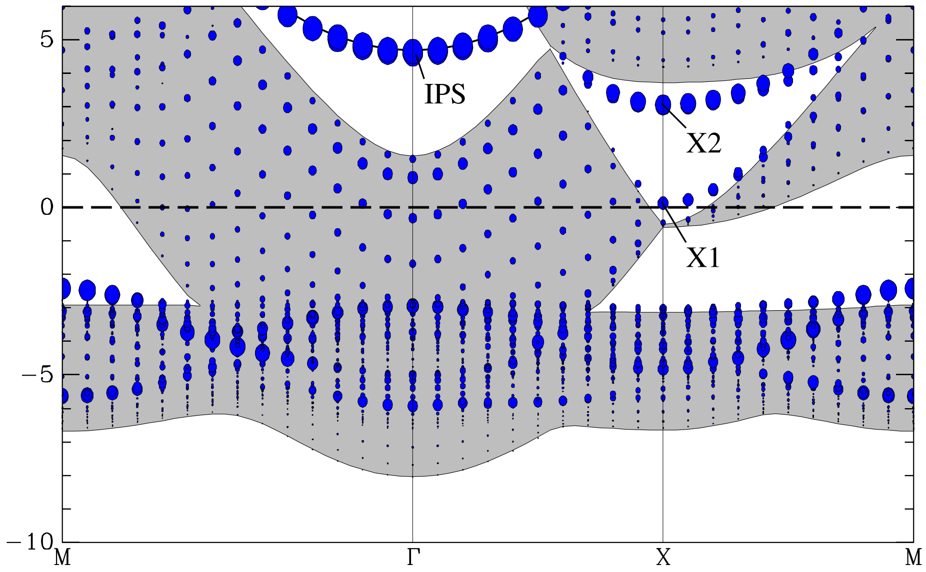


Figure 5.12: Calculated band structure for the Ag(100) surface (modeled with a 9 layer slab), compared with the (100) projection of the bulk energy bands (shaded area). The radius of the filled circles is proportional to the probability of the state to be outside the outermost ionic layer. The thick dashed line indicates the Fermi energy, which is aligned to zero. The thick continuous line (which is partially hidden by the IPS band) indicates the lower boundary of the continuum of free-electron states in the vacuum. Notice the unoccupied surface states, which are indicated as X1, X2 and IPS.

5.5 Ag bulk and (100) surface

We summarize in this section the main features of the theoretical and experimental electronic structure of the Ag(100) surface. In Fig. 5.12 we report the calculated LDA band structure for a relaxed 9 layers Ag(100) slab and compare it with the projected bulk band structure. We can easily recognize the localized 4*d* energy band which lies below eF, and the nearly free-electron like 5*sp* band, which is only partly occupied. The ion cores produce a rather strong effective periodic potential on the 5*sp* electrons, which leads to the opening of a large "stomach gap" around the center of the SBZ. The calculated work function for the slab is 4.68 eV, which is in agreement with previous experimental [97] and theoretical [86] studies. Therefore, the vacuum level at the Γ point lies approximately in the middle of the above mentioned gap.

We notice that just below the vacuum level there is an image-potential surface state which is well known both experimentally [97] and theoretically [95, 101]. In our calculations this surface state has a binding energy (0.05 eV) which is much smaller than the experimentally measured value ($\sim$ 0.5 eV). This is

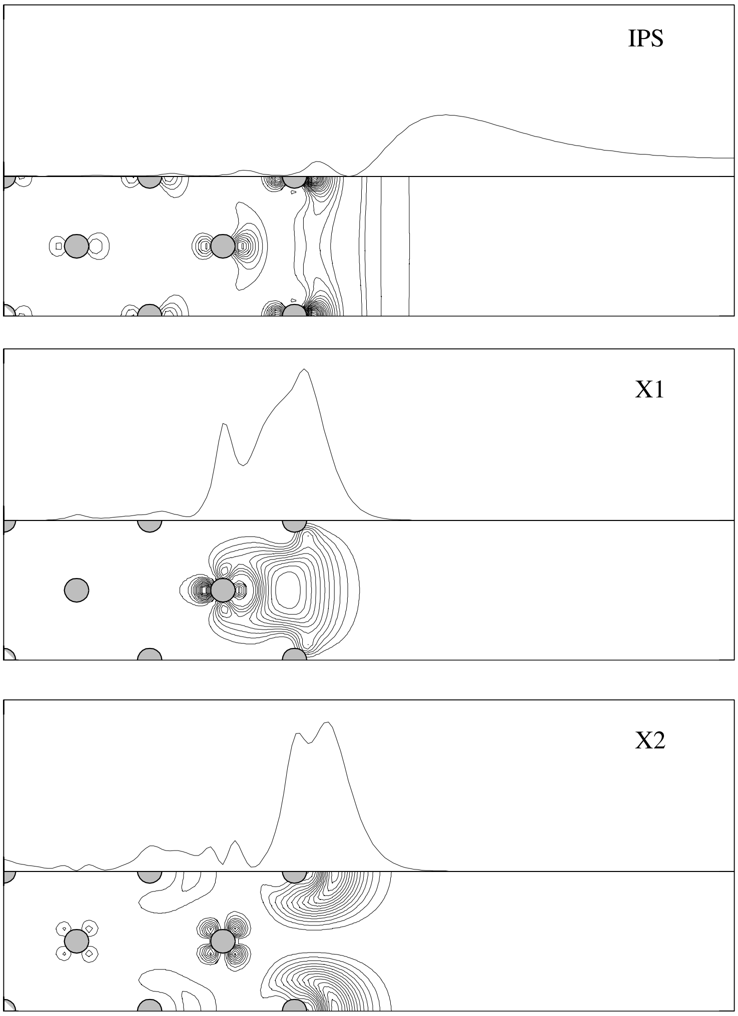


Figure 5.13: Charge density for the unoccupied surface states indicated in Fig.5.12 as IPS, X1 and X2. Ag atoms are represented by grey filled circles. Planar averages and contour plots are shown in the upper and lower part of each panel, respectively. We will see in following sections that X1 plays a role in the build-up of the surface dipole due to the interaction with MgO films, X2 is little affected by the presence of the overlayer while the image state disappears and is replaced by the IPS of the MgO/Ag(100) surface.

mainly due to the wrong treatment of exchange and correlation terms outside the metal in the LDA approximation, as discussed in Section 5.4.3.

We show in Fig. 5.13 the wavefunctions corresponding to the unoccupied localized surface states indicated in the surface band structure.

5.6 Supported MgO/Ag(001) films

Thanks to the favorable lattice parameter matching (3 % misfit between bulk Ag and bulk MgO) and to the almost ideal ionic character of magnesium oxide, the Ag/MgO interface is a very attractive model system for the study of the interaction between simple oxides and metals. The recent identification [69] of optimal conditions for epitaxial growth has allowed for an accurate experimental characterization of thin MgO films deposited on Ag(100). Theoretical studies and the results of DFT calculations corroborated those findings, and many issues concerning this system are now clear, while other remain controversial.

In the following we briefly review the existing experimental and theoretical literature, evidencing what has been done and what is still unclear, and then we will discuss in detail the results of the present work.

5.6.1 Review of existing literature

The structural properties of thin MgO films deposited on Ag(100) have been thoroughly investigated by means of STM [2, 65, 66], LEED [66], RHEED [70] and other electron diffraction techniques [72]. All studies report evidence for high-quality defect-free stoichiometric epitaxial films [2], whose in-plane lattice constants match the Ag substrate in the 1-5 ML range [70, 65], and gradually recover the MgO bulk lattice parameter for increasing thickness. Theoretical calculations for the MgO/Ag(100) interface [73] and adsorbed thin films [71, 70, 65] support the picture of ideal commensurate overlayers with a marked site preference for the O adsorption on top of Ag atoms. The interfacial distance has been carefully measured, and found to be between 2.4 and 2.5 A, depending very weakly on film thickness [72] and agrees very well with theoretical results. Theory further predicts a surface corrugation of ~ 0.1 A which, in the case of one monolayer, is larger than the reported value at the MgO(100) single crystal surface (0.04 A, see Section 5.3). This corrugation has been shown to correlate quite strongly to the magnitude of the dipole at the interface [71], which is responsible for an important lowering of the calculated workfunction.

From these studies the MgO/Ag(100) system is of a nearly ideal interface only slightly distorted by atomic relaxation, with crystalline 1-5 ML MgO(100) films growing commensurate with the substrate. This structural simplicity makes the accurate characterization of this system particularly interesting from a fundamental point of view, and has stimulated many electronic structure investigations. The occupied electronic states of 1 ML MgO/Ag(100) have been studied by means of UPS and a relatively strong interaction between oxygen $2p$ states and the silver $5sp$ band has been reported [68], while the Ag_{4d}-O_{2p} inter-

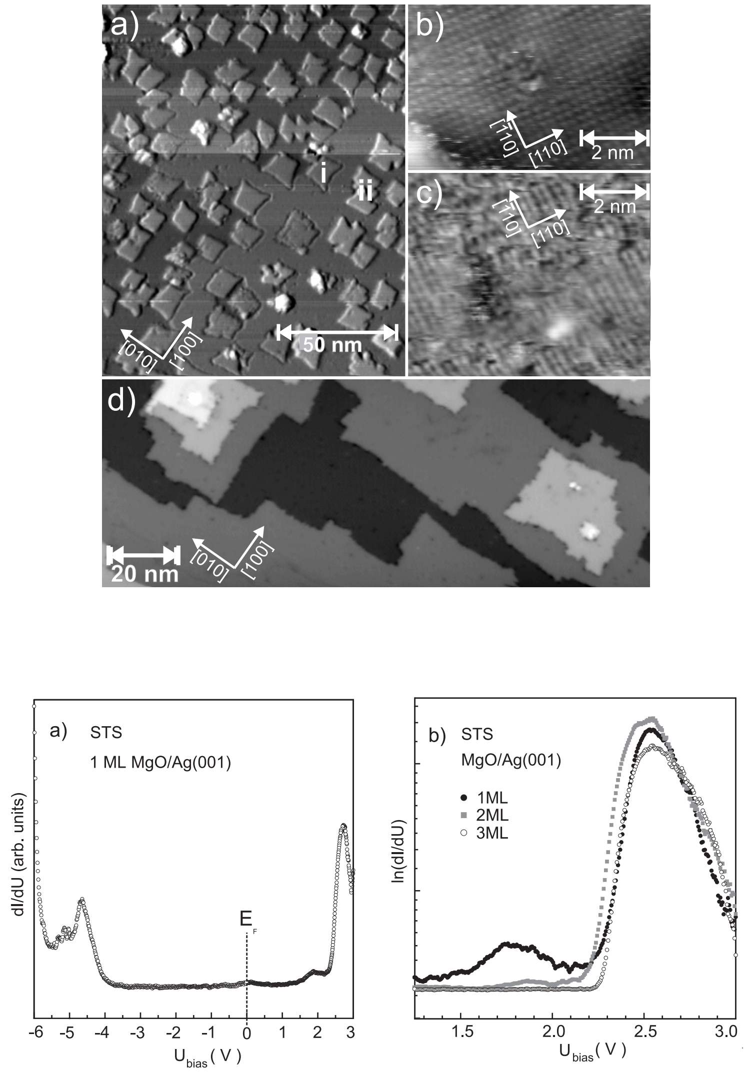


Figure 5.14: (Top) STM images at U bias of (a) 0.3 ML MgO/Ag(100), U= 5.0 V; (b) atomic resolution *through* an MgO island, U = 30 mV; (c) atomic resolution of the MgO layer (one type of ion is resolved), U=2.5 V; (d) 2.0 ML MgO/Ag(100), U=3.0 V. (Bottom) STS of ultrathin MgO/Ag(100) films: (a) tip placed above a 1 ML MgO island; (b) film thickness dependent dI/dU spectra. Data from [2].

action was found to be weak. Another UPS study [70] supports these findings, while the unoccupied states are investigated in the same work by EELS.

Finally, S. Schintke et al. [2] could obtain with STM and STS techniques very accurate information about the *local* surface electronic structure as a function of MgO film thickness. These authors found that already at one monolayer thickness the STS spectra show the typical features of the MgO(100) surface electronic structure, i.e. a spectral region about 6 eV wide which is almost flat, except for a small structure at 1.7 eV positive bias, which disappears when going to 2 and 3 ML (see Fig. 5.14). These observations were supported by our DFT calculations for the layer-resolved surface local density of states (LDOS), which showed a very rapid exponential damping of the LDOS in the surface gap as the MgO film thickness was increased from zero to three atomic layers.

In spite of this considerable research effort, many fundamental aspects remain unexplained. First of all, the striking independence of the measured energy position for the surface conduction band as a function of MgO film thickness was not accounted for by our calculations. Secondly, the feature at 1.7 eV observed by STS was tentatively assigned to an interface electronic state, but its physical origin could not be ascertained. More generally, surfaces of oxides and wide band gap insulators are not very well known, and as a consequence the experimental data for the supported thin films are often interpreted by means of qualitative macroscopic models, while a deeper understanding of the undergoing mechanisms is still lacking.

These open questions motivated us to pursue our theoretical studies, which eventually allowed for a more complete picture to emerge.

5.6.2 1-3 ML MgO/Ag(100) LDOS

We report in Fig. 5.15 the calculated surface LDOS corresponding to the 1-3 ML MgO/Ag(100) systems, compared with the LDOS for the clean Ag(100) (bottom) and MgO(100) (top) surface. The atomic positions for the MgO and the two outermost Ag layers have been fully relaxed, while the other Ag atoms are fixed in their ideal bulk positions. The in-plane lattice constant is taken as the calculated equilibrium lattice parameter for Ag bulk. The LDOS has been integrated in each case in the vacuum region starting from a fixed distance of 0.8 a.u. above the outermost atomic layer, in order to simulate the signal that the STS tip is most likely to detect during measurements. By comparing the calculated curves with the experimental data of Fig. 5.14 we can notice many common features:

- there are two sharp structures at negative and positive energy, which do not disappear for increasing MgO thickness;
- as the film thickness increases, the signal in the energy region between those structures is rapidly damped out, and only the typical features of the MgO surface remain present;

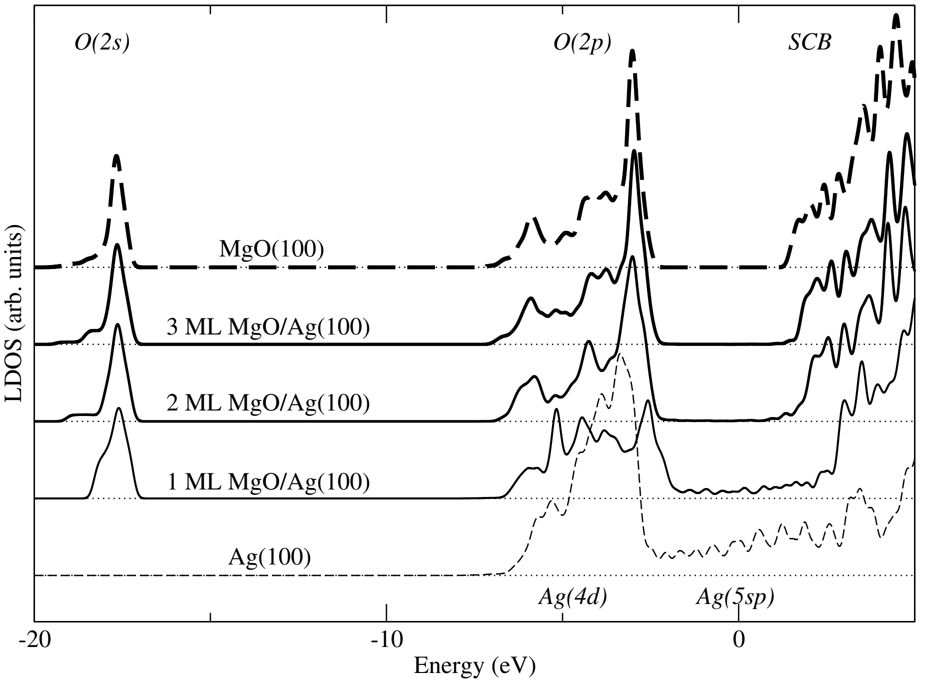


Figure 5.15: Surface LDOS for clean Ag(100) (thin dashed line), 1-3 ML MgO/Ag(100) (increasingly thick solid lines), clean MgO(100) (thick dashed line). All curves have been calculated by sampling the SBZ with a 16×16 uniform Monkhorst-Pack grid, and by integrating the charge density in xy planes and for z in the range between 0.8 a.u. above the outermost atomic layer and the center of the vacuum region in the periodic slab geometry. A Gaussian broadening of 0.15 eV has been used, which is a small value and gives good energy resolution but results in a somewhat wobbly curve for the Ag(100) $5sp$ band.

- when going from two to three ML the surface electronic structure does not change appreciably, and is already very close to that corresponding to the pure MgO(100) surface.

Thus the important result that three layers of MgO are sufficient to develop an electronic structure that is nearly indistinguishable from that of the single-crystal MgO surface is well reproduced in our calculations. However, the agreement of our calculations with experimental seems worse for the 1 ML case. In fact, Fig. 5.15 shows that the energy position of the calculated conduction band for 1 ML MgO/Ag is substantially higher in energy than in the case of thicker supported films. The valence band edge is also slightly higher in energy and its LDOS appears different in shape, but the shift of the conduction band dominates, and this leads to an important increase of the calculated surface electronic gap, which is in clear contrast with experimental observations.

In the next section we will investigate the reason for this discrepancy by analyzing in deeper detail the surface band structure for these three systems.

5.6.3 k-resolved LDOS

We report in Fig. 5.16 the complete electronic structure for the 1-3 ML MgO/Ag(100) systems calculated along the high-symmetry directions in the surface Brillouin zone. The bulk projected band structure of Ag is indicated by the shaded areas, and is aligned with respect to the Fermi level. Each energy level is indicated with a filled circle whose radius is proportional to the probability of the state to be in the vacuum region, defined as in the above discussion of the LDOS. Thus, these representations can be interpreted as the k-resolved analysis of the LDOS shown in Fig. 5.15, and allow for a detailed investigation of the main features discussed above.

The states close to Γ at the valence-band top of MgO(100) films have a predominant oxygen $2p_{xy}$ character, and interact very weakly with the substrate bands, as we can see from the contours in Fig. 5.18. In the single monolayer case the valence-band top is at the M point of the SBZ and shows a significant (repulsive) interaction with the underlying Ag $4d$ states, but remains well isolated in energy from the substrate energy bands. As a consequence, the surface band edge appears in all cases as a sharp feature in the LDOS, and its energy can be defined with high accuracy, since the distinction from metal-induced states is obvious in all cases.

Concerning the conduction band, the state corresponding to the sharp feature in the positive energy region of the LDOS is located in all cases inside a "stomach gap" of the bulk-projected electronic structure of Ag. Thus, this state can not penetrate into the metal and is constrained to the surface, as we can see from the charge density plots in Fig. 5.17. Moreover, since this state is not resonant with the Ag $5sp$ band continuum, its energy is again well defined in each case. Looking at the depicted charge density diagrams we can easily see that the shape and the physical nature of this state is in all cases very similar to that of the surface conduction-band state calculated for MgO(100) (see

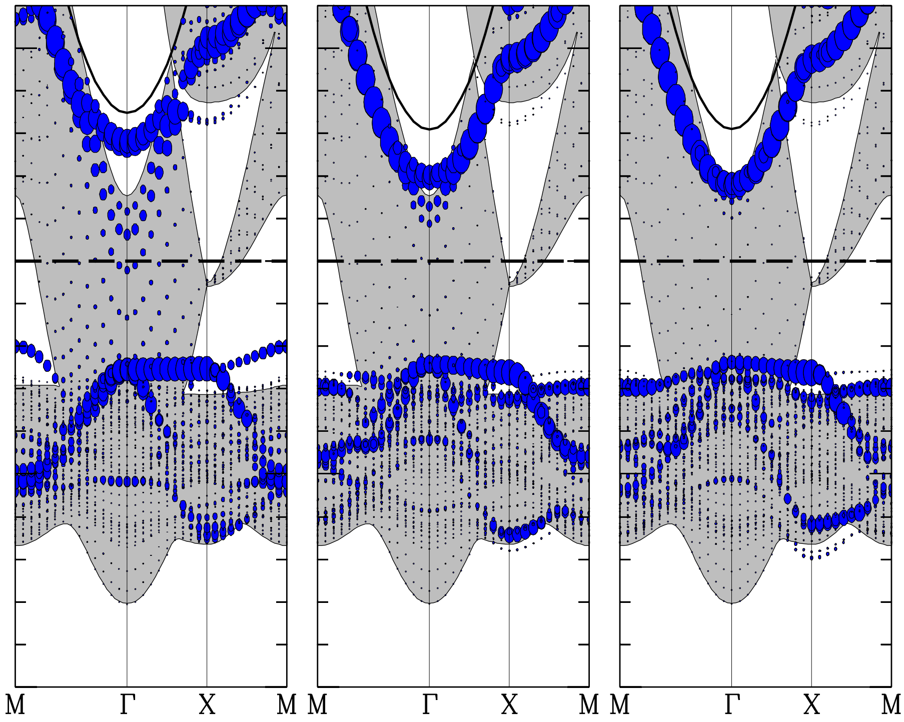


Figure 5.16: Calculated k-resolved LDOS of the MgO/Ag(100) system for MgO film thicknesses of 1 ML (left), 2 ML (center) and 3 ML (right). The Fermi level (broken horizontal line) is set as the zero of the vertical scale. Dashed areas indicate the projected band structure of bulk Ag. The area above the thick continuous curve is the range of the vacuum states, starting at $E_{vac} + k^2/2$ with k being the two-dimensional wavevector parallel to the surface. Notice the surface state at M for the monolayer case, which determines the surface valence-band top (as we have seen for the unsupported 1 ML film). The intensity of the MIGS decreases very quickly, and they are hardly visible for a 3 ML thick film.

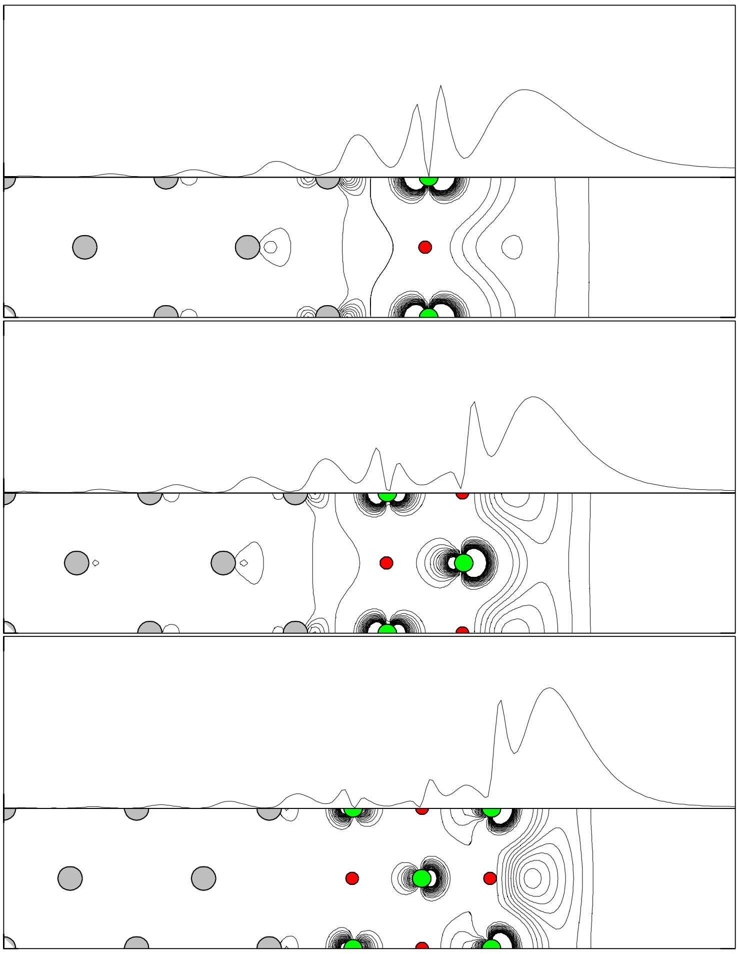


Figure 5.17: Charge density of the surface conduction-band edge state for the 1-3 ML MgO/Ag(100) systems. The charge density for the 3 ML case is practically identical to that of the surface state of pure MgO(100) (see Fig. 5.4).

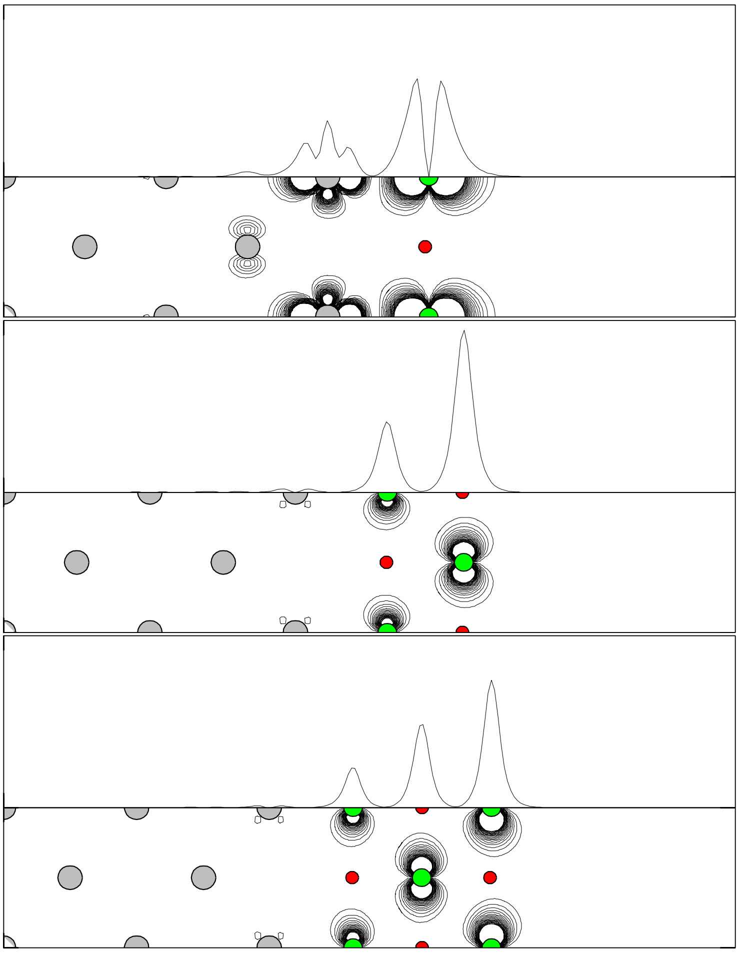


Figure 5.18: Highest surface valence state (as defined in the text) for 1-3 ML MgO films deposited on Ag(100). The valence band top in the 1 ML case lies at the M point of the SBZ, and has oxygen $2p_z$ character, while for thicker films it lies at Γ and has oxygen $2p_{xy}$ character.

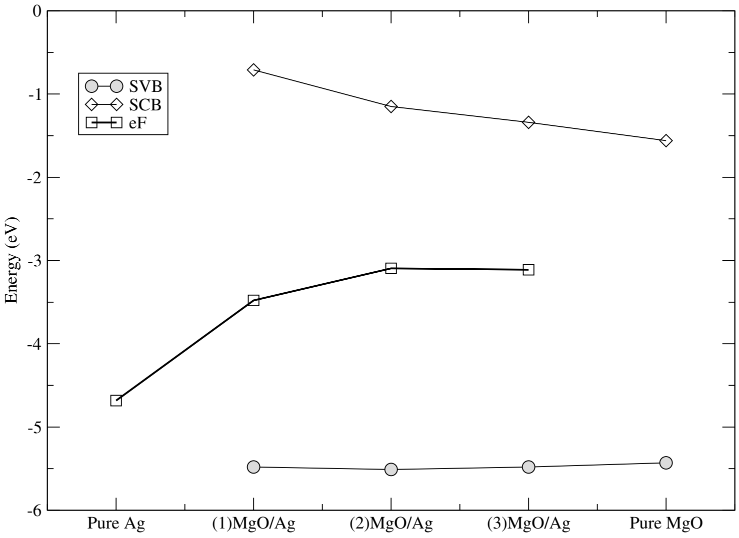


Figure 5.19: Surface conduction band (SCB), valence band (SVB) and Fermi level (eF) as a function of MgO layer thickness. The zero of the energy is set at the vacuum level for all systems. The calculated Fermi level of the pure Ag surface and the calculated SCB and SVB for the MgO(100) surface are also reported for comparison.

Fig. 5.4).

In summary, due to the particular features of the band structure of the MgO/Ag(100) system, we can always distinguish without ambiguity a surface conduction band edge (SCB) and a valence band edge (SVB), which can be directly related to the corresponding features of the pure MgO(100) surface. We further define the surface gap (SG) to be $E_{SG} = E_{SCB} - E_{SVB}$, and the MIGS will be referred to in the following as the electronic states with energy comprised between SVB and SCB. This analysis corroborates the interpretation of the experimental features in Fig. 5.14 as the MgO(100) surface valence and conduction bands, which was proposed in our previous work [2].

In order to better understand the influence of the substrate on these features it is useful to represent the corresponding energies on an *absolute* scale, where the vacuum level is taken as the energy zero. With this choice, we can directly compare our results for the (1-3)MgO/Ag(100) system with our findings for the pure MgO(100) surface, as shown in Fig. 5.19.

The SVB energy level changes little upon reducing the film thickness, while, the SCB undergoes a significant positive energy shift, which becomes increasingly large as the thickness is reduced to one monolayer. This effect is the main reason for the increase in the calculated surface gap. We will discuss in the following sections this feature, which is at variance with the experimental data, by using the arguments about image-potential states developed in Section 5.4 in the context of the MgO(100) surface.

5.6.4 Work function

We want now to discuss the variation of the Fermi level (i.e. the adsorbate-induced shift in the metal work function) as a function of the MgO film thickness. The work function is determined by the surface dipole, which is changed by the presence of the MgO film. In a rigid ion model, the inward rumpling of Mg would induce a work function increase, but this effect is compensated by a depletion of electrons in the $2p_z$ orbitals of oxygen atoms, which accumulates in the hollow sites of the Ag surface, thus reversing the sign of the work function shift. We show in Fig. 5.20 the changes in the self-consistent charge density which are due to the interaction with the metal substrate in the 1 ML case. Interestingly, the pattern of the accumulated charge corresponds approximately to the localized X1 state of the Ag(100) surface, which is empty in the case of the clean surface. Our results suggest that the presence of the positively charged Mg ions in the hollow site lowers the energy of the X1 state, thus restoring the degeneracy of the Ag *sp* band at the X point, which was broken by the presence of the metal-vacuum interface (see the bare Ag(100) band structure in Fig. 5.12). In our calculations we found that the charge transfer due to the oxide-metal interaction is well localized at the interface, and the calculated work function does not appreciably change when going from two to three MgO layers, saturating at 3.1 eV, while the calculated value for 1 ML is significantly higher (3.48 eV). The trend of the calculated values of the work function shows fair agreement with experimental data as a function of coverage, which are shown

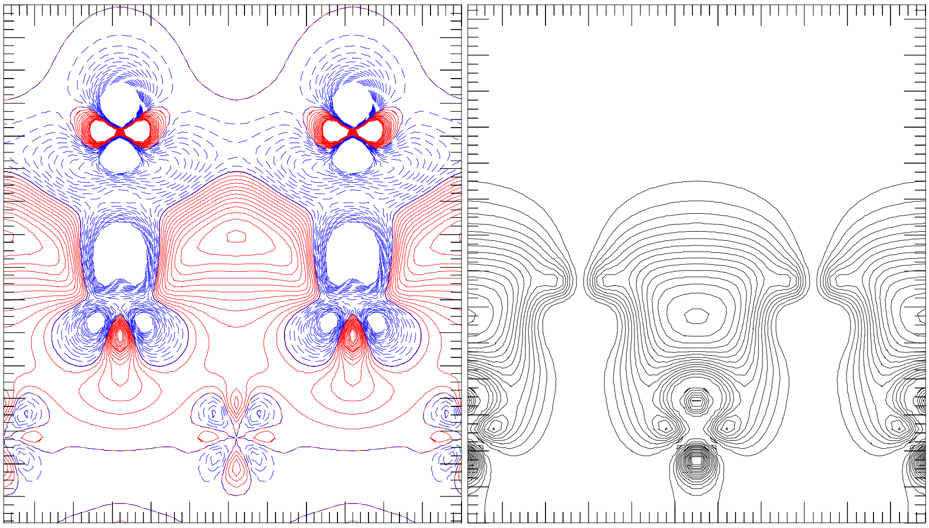


Figure 5.20: Left panel: charge density redistribution for the 1 ML MgO/Ag(100) system (red solid lines = accumulation, blue dashed lines = depletion). Right panel: charge density of the X1 unoccupied surface state of clean Ag(100).

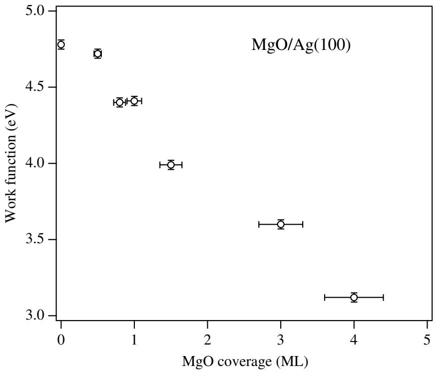


Figure 5.21: Experimental (UPS) data for the work function of the MgO/Ag(100) system as a function of nominal coverage (W.-D. Schneider, private communication).

in Fig. 5.21.

To investigate the dependence of the WF on surface rumpling, we calculated the WF for a 1 ML MgO/Ag(100) system without rumpling, i.e. with atomic positions for the oxide layer assumed to be the average value between the Mg and O vertical coordinates in the fully relaxed configuration. The workfunction in this case was reduced to 2.87 eV. This dependence of the WF on the atomic relaxation at the interface can explain the WF reduction between 1 ML and thicker MgO films, since in the latter case the presence of additional ionic material reduces substantially the relative displacement of interfacial anions and cations.

5.6.5 Surface conduction band

The following discussion will focus on the unoccupied MgO(100) surface state at Γ, which we have shown to be the main factor which determines the variation of the surface gap as a function of film thickness in our calculations. We want here to understand on the basis of simple arguments the description of the system provided by LDA, and why it fails to reproduce the experiment.

In the case of MgO(100) we have seen that the surface conduction band is reproduced in the LDA approximation as a state bound to an effective potential well close to the surface. In a rough approximation, we can assume that the shape of this potential well does not depend too much on the thickness of the MgO film. From this point of view, we can qualitatively look at the MgO film as a spacer layer which separates the potential well from the Ag surface. By reducing the MgO film thickness, this potential well is closer to the metal surface, and the corresponding bound state can interact with the silver electronic states which are almost degenerate in energy. Our calculations indicate that (i) this interaction has repulsive character and (ii) it is rather strong, since it leads to a positive shift as large as 0.6 eV when going from 3 to 1 ML.

We can justify point (i) by the following argument. In our calculations the energy of the MgO film surface state is slightly higher than the top of the $5sp$ band at Γ, i.e. it lies in the Ag(100) surface bandgap (as we have seen in the previous section), close to the bottom. The wavefunctions of the top of the Ag $5sp$ band have *antibonding* character and they "repel" the approaching electronic state bound within the surface potential well, thus leading to a positive shift of its energy.

Concerning (ii), the magnitude of this shift depends mainly on the overlap between the wavefunction of the surface state and the metal Bloch waves. Thus, the energy shift depends critically on the penetration length of the surface state wavefunction in the region inside the material. In the previous sections we have concluded that LDA does not describe properly this surface state. In particular, we have deduced that this state is strongly overbound, its wavefunction is too close to the surface and its penetration length inside the material is overestimated. All these effects lead to a spurious enhancement of the sensitivity of this state to the details of the bulk material lying underneath the surface atomic layer. For these reasons, we must conclude that the surface gap increase

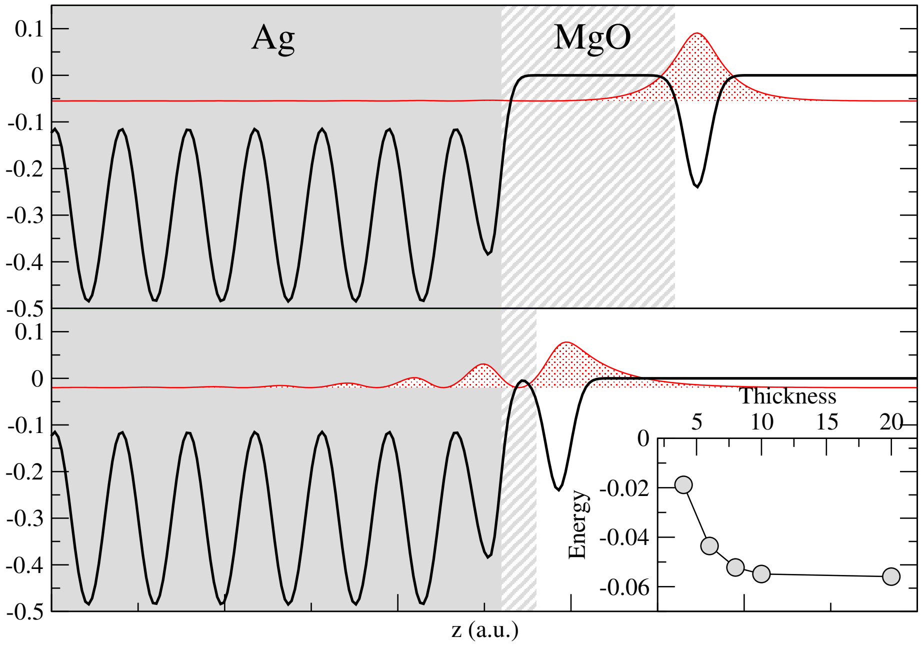


Figure 5.22: Simple 1D model for the LDA description of the surface unoccupied band in the MgO/Ag(100) system (see text for details). Upper panel: MgO film of 3 atomic layers (12 a.u.); lower panel: one monolayer (4 a.u.); inset: variation of the energy of the surface bound state as a function of MgO thickness.

for decreasing deposited MgO film thickness is unphysically enhanced by the LDA approximation, and we should not expect quantitative agreement with experimental results on this point.

To check the validity of our analysis, and better understand the failure of LDA, we study a 1D model of the MgO/Ag system which has the following features:

- the potential in bulk Ag is a simple cosine function, adjusted to reproduce the correct gap in the 5*sp* band at Γ, and aligned with respect to the vacuum according to the LDA results for the MgO/Ag(100) workfunction;
- the interface Ag layer is modified to reproduce correctly the surface unoccupied resonance just below the gap [94, 95];
- the potential in the MgO film is set to the vacuum level, consistent with the calculated MgO(100) electron affinity;
- the surface potential is modeled by a Gaussian well, adjusted to reproduce the calculated binding energy in the limit of high thickness, and its shape is kept constant for varying thickness (i.e. only the position with respect to the Ag surface is changed).

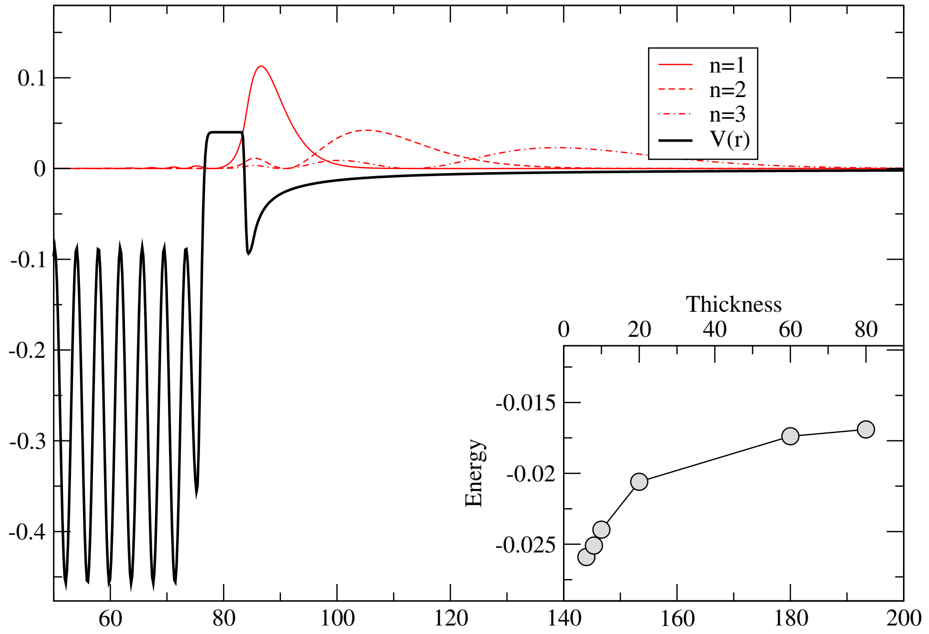


Figure 5.23: Improved model for the Ag/MgO/vacuum double interface, keeping into account the correct (classical) dielectric response of the combined metal/oxide system. We show the effective 1D potential (thick black line), the charge densities of the first three image states (thin red lines) and the evolution of the energy eigenvalue of the first IPS as a function of film thickness (inset, energy zero set at the vacuum level). All energies are in Hartree units.

We show in Fig. 5.22 the potential for two different thicnesses (4 a.u. and 12 a.u.) together with the results for the wavefunctions and energy eigenvalues. The behaviour of LDA is fully reproduced by this model, and therefore we ascribe the spurious LDA gap opening effect to the interaction between the MgO surface conduction state and the Ag 5sp band, as discussed above.

In order to understand whether the gap opening should be considered a physical effect or an LDA artifact we improved the above model as follows:

- the potential in the film is set to the MgO EA deduced from the experimental results (1 eV);
- the potential in vacuum is set to the classical response to an external electron of a homogeneous dielectric film supported by a perfect metal.

The solution to the boundary value electrostatics problem for the metal/thin film system was first obtained by Cole [93] and has similar form to the solution we calculated analytically for the thin unsupported film:

$$V(z) = -\frac{\beta}{4z} + \frac{1-\beta^2}{4\beta} \sum_{n=1}^{\infty} \frac{(-\beta)^n}{z+nt} \tag{5.7}$$

We show in Fig. 5.23 the results for the potential, wavefunctions and energy eigenvalues as a function of thickness. In striking contrast with the LDA results, the image states are now shifted to lower energies for thinner films, as a consequence of the dielectric response of the system.

We can now use this result to interpret the STS spectra in Fig. 5.14, for which the LDA description is clearly unsatisfactory. Analysis of the spectra reveals that they are not completely overlapping, i.e. the 2 ML peak is lower in energy than the 3 ML peak, and this is in agreement with our model. However, the 1 ML peak appears at the same energy as the 3 ML peak, and not lower than the 2 ML one as predicted. This is due to the fact that all curves are aligned with respect to the Fermi level of the metal, and not to the vacuum level. Taking into account the calculated variation of the work function and aligning the peaks in Fig. 5.14 with respect to the vacuum level, the predicted energy ordering ($E_1 < E_2 < E_3$) is fully reproduced, which gives an important experimental verification of our improved model.

5.6.6 Comparison with the experiments

We will now discuss two important experimental issues, namely (i) what is the precise origin of the peak at 1.7 eV in Fig. 5.14 and (ii) which ion, Mg or O, appears as a protrusion in the atomically resolved STM images at 2.5 eV positive bias. The calculated k-resolved LDOS indicate as a candidate for (i) the Ag states situated at the top of the $5sp$ band at the Γ point. These states are close in energy to the MgO surface conduction band, and as a consequence their exponential tails penetrate in the MgO film. However, since they lie in a forbidden energy range for propagation in MgO bulk, they are rapidly damped for increasing oxide film thickness, as we can see in Fig. 5.15, and this is in good agreement with experimental measures. Therefore, we propose an interpretation for the feature at 1.7 eV as MIGS, and not as a localized interface state as was tentatively reported in our previous work, since the existence of such a state is not predicted by our calculations.

Concerning (ii), we argue that the protrusions in STM images correspond to Mg ions. The STM tip is expected to follow the charge corrugation of the surface LDOS [110], and at a bias of 2.5 eV the main signal should come from the MgO(100) surface state, which has local maxima on Mg (see Fig. 5.4). In order to directly appreciate this fact, we have calculated the "simulated STM images", according to the modified Tersoff-Hamann theory which was already used in [111]. Namely, we compute the virtual charge density corresponding to the sum of the unoccupied wavefunctions having eigenvalue comprised between

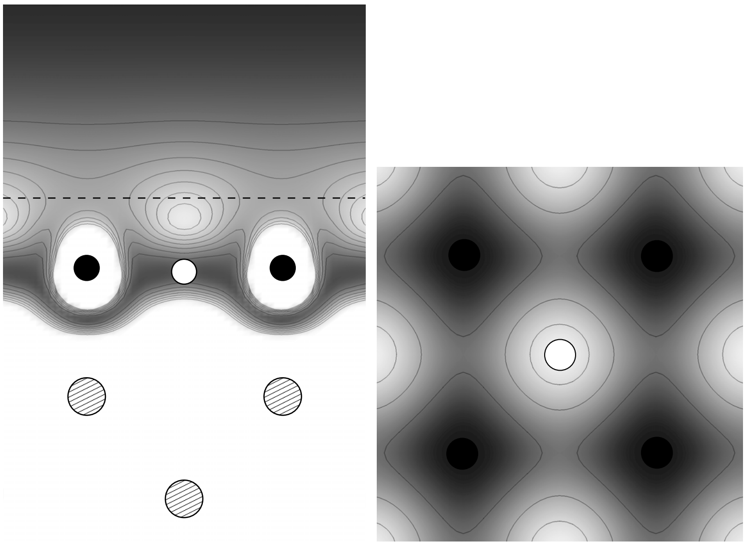


Figure 5.24: Energy-resolved charge density contour plots (see text for explanation) for the 1 ML MgO/Ag(100) system. Left panel: vertical view, the dashed line shows the position of the plane where the horizontal view given on the right-hand side has been taken. Silver atoms are indicated by large striped circles, solid and empty circles indicate oxygen and magnesium atoms, respectively.

E_F and a positive bias of 3.0 eV [7]:

$$\tilde{n}(\mathbf{r}) = \int_{E_F}^{E_{bias}} |\psi_\epsilon(\mathbf{r})|^2 d\epsilon \tag{5.8}$$

and we depict in Fig. 5.24 the contour plots of the density in the xz and xy planes. It is evident from the figure that at this bias the dominant contribution in the vacuum region is given by the SCB edge (the lobes above Mg atoms are very similar to the corresponding feature in the MgO(100) surface state). This leads to a protrusion in the charge density profiles above the Mg ions, consistent with the above discussion. Furthermore, our calculations predict a contrast reversal, i.e. for negative biases the tunnelling from the anion p states should dominate, in agreement with the observations for the NaCl/Al system [111].

The Ag region is completely white in the figure, since between 0 and 3 eV there is a substantial contribution from the Ag $5sp$ bands. Interestingly, there is a rather strong maximum in the region corresponding to the oxygen $2p_z$ orbitals.

---

[7] The experimental bias for atomically resolved images was 2.5 eV, and this corresponded to the main peak associated to the surface conduction band (see STS spectra in Fig. 5.14). In our calculations, due to the spurious LDA gap opening effect discussed in the previous section, the SCB appears 0.5 eV higher than in the experiments (w.r.t. Fermi level), and for a comparison we consider more reasonable to choose a correspondingly higher bias of 3.0 eV.

This means that an appreciable density of states with oxygen $2p$ character is pushed *above* the Fermi level upon interaction with Ag, which is in agreement with the experimental results of [68]. This confirms the induced depletion of electrons in the oxygen $2p_z$ orbitals that we already observed in the $\Delta\rho$ study. We notice that this is also consistent with the presence of MIGS in the gap of the insulating film, since these states are expected to take their weight primarily from the occupied bands which are nearest in energy, i.e. the valence band for $\epsilon_i < E_F$ [107].

5.6.7 Summary and outlook

We presented a detailed theoretical study of MgO(100) thin films deposited on an Ag(100) substrate. Particular emphasis was given to the variation of the surface electronic properties as a function of film thickness, which were investigated by separating the effects due to the reduced dimensionality of the oxide film from those arising from the interaction with the metal substrate. Our results allow for an unambiguous interpretation of the experimental data. In particular, we have shown that (i) the conduction band of the $(\mathrm{MgO})_n/\mathrm{Ag}(100)$ system can be interpreted as an image-potential induced surface state; (ii) the signal in the gap region can be ascribed to MIGS; (iii) the atomic species imaged as bright spots by STM are likely to be Mg. Furthermore, our findings are in substantial agreement with experimental measures, and provide a simple physical model for the (ideal) interface Ag/MgO/vacuum.

From the theoretical point of view we pointed out the merits but also the difficulties of DFT-LDA when discussing electronic properties, in particular the unoccupied bands. We have proposed a dielectric continuum model, in order to take into account the effects of the dynamical response on the electronic properties of the system. While these models are crude, they show that the energy gap variation may even be reversed with respect to the LDA predictions. One should then be particularly careful when analyzing the electronic structure of multi-interface systems (metal-insulator-vacuum), and not to overemphasize the meaning of the calculated trends.

Chapter 6

Conclusions

We have studied the adsorption of C_{60} molecules and MgO thin films on low-index metallic single-crystal surfaces. The study has been carried out from first-principles, in the framework of density functional theory within the local density approximation. For the minimization of the electronic structure and atomic relaxations we used a recently developed scheme [38], which extends the Car-Parrinello method to the treatment of metallic systems. In Chapter 3, the difficulties of this new scheme have been analyzed in depth, and possible improvements were investigated; in particular, a Lagrangian formalism was introduced and tested. In Chapter 4, we have presented our results on the electronic and atomic structure of the C_{60}/Al(111) system at one monolayer coverage; adsorption geometries with one or three C_{60} molecules per surface unit cell, deposited on a regular or defective substrate, were considered. In Chapter 5, we have reported our findings on the structural and electronic properties of the MgO/Ag(100) system for a film thickness ranging from one to three atomic layers; the properties of the MgO(100) single-crystal surface were also discussed in detail.

The main results achieved are:

- **C_{60}/Al(111): covalent bonding**

  We find that the interaction with the Al substrate induces appreciable structural distortions of the adsorbed C_{60} molecules. The results for the charge density indicates partial sp^3 rehybridization of the carbon atoms which are closest to the surface Al atoms. Both results are in good agreement with experimental findings, and with the predominantly covalent bonding picture which was previously reported. Furthermore, we find that this charge redistribution is responsible for a structural instability which is predicted for the most symmetric adsorption geometry. We show that this instability can be relieved by the creation of a vacancy in the substrate, which allows for an optimal C_{60}-Al bonding with a much higher adsorption energy. This strong preference for a vacancy site could explain the occasional nucleation of C_{60} islands far away from step edges, which was observed by STM and tentatively ascribed to the presence of isolated defects in the substrate.

- **C_{60}/Al(111): mechanisms for the (6×6) reconstruction**

  We find that the covalent nature of the C_{60}-metal interaction plays a dominant role in the reconstruction mechanism, rather than the compressional stress in the overlayer. Our results indicate that the energy gain resulting from an improved C_{60}-Al coordination is large enough to overcome the high energy cost of creating point defects in the Al substrate. This yields a reconstruction mechanism where Al vacancies are created at the adsorption sites together with adatoms which move to the interstitial regions underneath the C_{60} overlayer. This mechanism is consistent with the experimental evidence of Al adatoms in the reconstructed structure, and with the observation that C_{60} molecules which protrude from the surface interact strongly with the underlying aluminum.

- **Structural properties of unsupported MgO(100) thin films**

  We find that thin MgO films have a substantially smaller in-plane equilibrium lattice constant than that of a bulk MgO crystal. This result is in qualitative agreement with a previous theoretical study, and explains the experimental observation that thin MgO films tend to grow commensurate with an Ag(100) substrate despite the appreciable lattice mismatch between the two bulk materials.

- **MgO/Ag(100): electronic properties**

  We find that the surface conduction band of the MgO/Ag(100) system has the character of an image state, and lies in an energy region where propagation into bulk Ag is forbidden. The states at the valence band edge of the MgO film interact only weakly with the substrate, and the corresponding energies do not show appreciable variation as a function of film thickness. The LDOS in the energy range between the above two features is dominated by the tails of metallic wavefunctions which tunnel across the MgO film. The latter are rapidly damped to zero for thicknesses larger than 2 ML. These results allow for the unambiguous interpretation of the features observed in the STS spectra, whose variation as a function of MgO thickness is qualitatively well reproduced by our calculated LDOS. Furthermore, by simulating the STM images at the experimental value of bias, we conclude that the Mg ions are imaged as protrusions in the atomically resolved topographs. Finally, we calculate the work function variation with film thickness and we find overall agreement with the experimental values as a function of nominal coverage; we point out the role of atomic relaxation in determining the magnitude of the surface dipole, thus improving previous theoretical estimates [71].

Our study has also provided some results that need further experimental investigations and/or theoretical study to be finalized:

- **Molecular dynamics for metallic systems**

  We find that the scheme proposed by Vandevondele and De Vita [38] (VDV) is efficient and reliable for both electronic and structural optimizations in a wide range of metallic and nonmetallic systems. The method however presents some difficulties when performing microcanonical molecular dynamics simulations: (i) the absence of a rigorous constant of motion and (ii) the breakdown of adiabaticity during long runs. We have addressed (i) by constructing a Lagrangian for the ionic, electronic and occupational degrees of freedom, where the fictitious mass terms are preconditioned according to the gradients of the total energy functional. The resulting equations of motion coincide in first order to those of the original scheme, and thus provide a sound mathematical foundation to this method. Concerning (ii), we show that the complete Eulero-Lagrange equations are able to improve dramatically the adiabatic regime in some selected cases (liquid Al and Si_2). The new formulation, however, is less robust than the original scheme, and thus it is not yet suitable for general applications. We have discussed possible future improvements of the method.

- **C_{60}/Al(111): adsorption sites**

  We investigated several adsorption sites for the C_{60} adsorption on the Al(111) surface, and we found that the most favorable one is not the symmetric *atop* one as it was previously speculated. We find three equilibrium configurations which are quite close in energy with each other, and this could explain the observed high mobility of adsorbed C_{60} molecules at room temperature. This mobility, which does not seem compatible with the reported strong covalent bonding, could be related to the existence of many bonding configurations which are close in energy. However, for a definitive answer, a theoretical study of the diffusion barriers and an experimental determination of the preferred adsorption sites are needed.

Finally, our work evidenced some new interesting problems and suggested directions for future activities:

- **Ionic films on metallic surfaces**

  The influence of a thin adsorbed film on the image-state spectrum of a metal is of outstanding scientific interest, and very accurate results have been obtained very recently e.g. for Ar/Cu(100) and C_{60}/Cu(111). Our results suggest that metal-supported ionic thin films like NaCl(100), LiF(100) and MgO(100) could have many features in common with the above systems. Thus, the theoretical tools (dielectric continuum models) and experimental techniques (two-photon photoelectron spectroscopy) that have been successfully used in the investigation of femtosecond electron dynamics at metallic surfaces could be equally useful for the study of ionic crystals, calling then for a unification of these research fields.

- **MgO/Ag(100): insulator at the ultrathin limit**

  The emerging field of insulating supported thin films calls for novel interpretations of some well-estabilished physical concepts. Strictly speaking, few oxide layers deposited on a metal surface are not insulating at all, since the tails of the Bloch waves of the metal in the vicinity of the Fermi level can tunnel across the film, and provide metallic character. The metallicity is further confirmed by the fact that current transport is possible, and even a 15-layer thick film does not show appreciable charging upon spectroscopical probing. So, "insulator" in this case means something different than usual, and is related to the dielectric layer capability of efficiently decoupling the surface region from metallic electronic states, within the energy range corresponding to the bulk band gap of the insulator. This is particularly relevant in the study of surface chemical reactions or supported nanostructures, since the direct contact with a bare metallic surface can strongly affect the electronic properties of an adsorbed molecule (e.g. the discrete molecular energy levels are broadened into resonances), with unwanted consequences on chemical reactivity and/or catalytic properties. From this point of view we find that a three-layer thick MgO film is *insulating* on Ag(100), since the surface electronic structure of the supported film is hardly distinguishable from that of a MgO single-crystal. Interestingly, two- and one-monolayer supported MgO films appear qualitatively very similar, apart from the appearence of a finite density of states in the surface gap, which we interpret as metal-induced gap states (MIGS), introduced long ago in the context of metal-semiconductor junctions. This concept applies equally well to the discussion of metal-insulator interfaces, consistent with the recent identification of MIGS in the LiCl/Cu(100) [114] system. This property of the insulating layer of decoupling the surface region from the metal bulk in a controllable, thickness-dependent manner opens new perspectives for the development of oxide heterostructure-based nanodevices and applications in magnetoelectronics.

In conclusion, this thesis demonstrated once more the usefulness of the first-principles approach in the characterization of realistic surfaces and nanostructures. We considered two complementary systems of current technological and fundamental interest, which provided different problems and needed different approaches from the theoretical and methodological point of view. In both cases we could achieve a satisfactory physical model, give original interpretations of experimental data, and formulate predictions that may be valuable for future work.

Appendix A

Derivation of the Eulero-Lagrange equations of motion

Our scope is to derive the Eulero-Lagrange (EL) equations of motion from the Lagrangian expression and from the holonomic constraints $\langle\psi_i|\psi_j\rangle = \delta_{ij}$ and $\sum_i f_i = N_e$ which enforce the orthonormality of states and the conservation of the number of electrons. For this we need to study the dependence of $\mathbf{f}$ and $\dot{\mathbf{f}}$ on the variables of motion ξ and $\dot{\xi}$. The equations of motion will be always expressed in the reference system $\mathcal{O}$ where the ξ and $\mathbf{f}$ matrices are diagonal.

We first investigate the dependence of the f matrix on the ξ_{ij} degrees of freedom. After a first-order variation $\delta\xi$, the matrix $\xi + \delta\xi$ will be in general no more diagonal in $\mathcal{O}$. We express the infinitesimal transformation that recovers the new diagonal reference system $\mathcal{O}'$ (to which primed quantities will refer) by a first-order antihermitian matrix W:

$$[(1 - W)(\xi + \delta\xi)(1 + W)]_{ij} = \xi'_{ij} = \delta_{ij}(\xi_i + \delta\xi'_i) \qquad \text{(A.1)}$$

W is well defined if the diagonal elements of ξ are non degenerate, with $W_{ij} = \delta\xi_{ij}/(\xi_j - \xi_i)$. From Eq. A.1 we obtain a relation for the $\delta\xi_{ij}$ which must also hold for the occupancies, since $[\mathbf{f}, \xi] = 0$:

$$\delta\xi_{ij} = \delta_{ij}\delta\xi'_i + [W, \xi]_{ij} \qquad \text{(A.2)}$$

$$\delta f_{ij} = \delta_{ij}\delta f'_i + [W, \mathbf{f}]_{ij} \qquad \text{(A.3)}$$

Substituting the W_{ij} expression into eq. [A.3] we obtain

$$\delta f_{ij} = \frac{f_j - f_i}{\xi_j - \xi_i}\,\delta\xi_{ij} \qquad i \neq j \qquad \text{(A.4)}$$

while for diagonal elements we have simply

$$\delta f_{ii} = \delta f'_i = [\frac{\partial f_{FD}(\xi)}{\partial \xi}]_{\xi=\xi_i}\, \delta\xi_{ii} \tag{A.5}$$

Inspection shows that these relations can be extended continuously to the general case of degenerate ξ_i. For this, we define the real symmetric matrix $\bar{f}_{ij} = G(\xi_i, \xi_j)$, through the two-variable function $G(x, y)$:

$$G(x,y) = \begin{cases} \frac{f_{FD}(x)-f_{FD}(y)}{x-y} & x \neq y \\ \frac{1}{2}[\frac{\partial f_{FD}(x)}{\partial x} + \frac{\partial f_{FD}(y)}{\partial y}] & x = y \end{cases} \tag{A.6}$$

Notice that the function $G(x, y)$ is continuous in $x = y$ and that its first-order partial derivatives are everywhere well defined. Also, since $f_{FD}(x)$ is monotonically decreasing, $\bar{f}_{ij} < 0\ \forall i, j$. Finally, this simple relation holds in the general case:

$$\delta f_{ij} = \bar{f}_{ij}\delta\xi_{ij} \qquad \forall i, j \tag{A.7}$$

We can thus think of each matrix element f_{ij} in the $\mathcal{O}$ reference frame as a function of the correspondent element ξ_{ij} which undergoes the infinitesimal transformation, so that

$$\frac{\partial f_{ij}}{\partial \xi_{ij}} = \bar{f}_{ij} \tag{A.8}$$

which also provides a straightforward way to express the time derivative of the $\mathbf{f}$ matrix in the $\mathcal{O}$ reference frame:

$$\dot{f}_{ij} = \bar{f}_{ij}\dot{\xi}_{ij} \tag{A.9}$$

With this definition the kinetic energy term associated to the ξ variables in the Lagrangian $\mathcal{L}$ reads

$$-\frac{1}{2}Q\mathrm{T}r(\dot{\mathbf{f}}\,\dot{\xi}) = -\frac{1}{2}Q\sum_{ij}\bar{f}_{ij}|\dot{\xi}_{ij}|^2 \tag{A.10}$$

where Q is a positive mass constant. The dependence of the matrix $\dot{\mathbf{f}}$ on the dynamical variables is slightly more complicated, because this matrix depends on both ξ and $\dot{\xi}$. We will again start by expressing the first-order variation of $\dot{\mathbf{f}}$ in the $\mathcal{O}$ reference frame upon first-order changes of ξ and $\dot{\xi}$. It is helpful to isolate in the variation the part due to the changes in the off-diagonal elements of ξ from all the rest:

$$\delta\dot{f}_{ij} = \delta^{offd}\dot{f}_{ij} + \delta^{other}\dot{f}_{ij} \tag{A.11}$$

We thus suppose a non zero variation $\delta\xi_{ij}$ of the offdiagonal elements ξ_{ij}. As before, provided that the ξ_{ii} are non degenerate, this determines the first-order W and the reference frame $\mathcal{O}'$ where the rotated matrix $\xi' = (1-W)(\xi+\delta\xi)(1+W)$ is diagonal. Therefore, the matrix $\dot{\mathbf{f}}$ after the variation can be computed in $\mathcal{O}'$ using eq.[A.9]

$$[(1-W)(\dot{\mathbf{f}}+\delta^{offd}\dot{\mathbf{f}})(1+W)]_{ij} = \bar{f}'_{ij}((1-W)\dot{\xi}(1+W))_{ij}$$

We remark that in the last equation $\bar{f}'_{ij} = \bar{f}_{ij}$, since the matrix $\bar{f}$ is built from the diagonal elements f_i and ξ_i, whose values are not changed by this transformation (since $\delta\xi_{ii} = 0$, and from eq. [A.5]). The final expression for the variation of $\dot{\mathbf{f}}$ due to a first-order rotation is simply written using commutators:

$$\delta^{offd}\dot{f}_{ij} = [W,\dot{\mathbf{f}}]_{ij} - [W,\dot{\xi}]_{ij}\bar{f}_{ij} \tag{A.12}$$

Substituting the explicit expression for W we obtain

$$\delta^{offd}\dot{f}_{ij} = \sum_{k\neq j}\mathcal{F}_{ijk}\dot{\xi}_{ik}\delta\xi_{kj} + \sum_{k\neq i}\mathcal{F}_{jik}\delta\xi_{ik}\dot{\xi}_{kj} \tag{A.13}$$

where the quantity $\mathcal{F}$ is defined by

$$\mathcal{F}_{ijk} = \begin{cases} \frac{\bar{f}_{ij}-\bar{f}_{ik}}{\xi_j-\xi_k} & \xi_j \neq \xi_k \\ \left[\frac{\partial G(\xi_i,\xi)}{\partial\xi}\right]_{\xi=\xi_j} & \xi_j = \xi_k \end{cases} \tag{A.14}$$

Inspection shows that with this definition the expression [A.13] extends correctly to the degenerate case in which $\xi_j = \xi_k$ for some indices j,k. Moreover $\mathcal{F}_{ijk}$ is invariant under any permutation of the three indices. The remaining part of the variation in eq. [A.11], taking into account the dependence of $\dot{\mathrm{f}}$ on the diagonal elements of ξ and on the $\dot{\xi}$ variables, is readily evaluated from eq. [A.9]:

$$\delta^{other}\dot{f}_{ij} = \sum_{l}\frac{\partial\bar{f}_{ij}}{\partial\xi_l}\dot{\xi}_{ij}\delta\xi_l + \bar{f}_{ij}\delta\dot{\xi}_{ij} \tag{A.15}$$

Putting together eqns. [A.12] and [A.15] we obtain a compact expression for the first-order variation of $\dot{\mathrm{f}}$:

$$\delta\dot{f}_{ij} = \sum_{k}\mathcal{F}_{ijk}(\dot{\xi}_{ik}\delta\xi_{kj} + \delta\xi_{ik}\dot{\xi}_{kj}) + \bar{f}_{ij}\delta\dot{\xi}_{ij} \tag{A.16}$$

which is valid for all i,j indices and will now be used to derive the explicit form of the EL equations of motion for the ξ degrees of freedom. The EL equations read:

$$\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\xi}_{ij}} - \frac{\partial\mathcal{L}}{\partial\xi_{ij}} = 0 \tag{A.17}$$

The partial derivatives that enter these equations are:

$$\begin{aligned} \frac{\partial\mathcal{L}}{\partial\dot{\xi}_{ij}} &= -Q\dot{f}_{ji} \\ \frac{\partial\mathcal{L}}{\partial\xi_{ij}} &= -\bar{f}_{ji}(\tilde{h}_{ji} - \xi_{ji} + \lambda\delta_{ji}) - \frac{1}{2}Q\frac{\partial\mathrm{Tr}(\dot{\mathrm{f}}\,\dot{\xi})}{\partial\xi_{ij}} \end{aligned} \tag{A.18}$$

where $\tilde{h}_{ji} = h_{ji} - \mu\langle\dot{\psi}_j|\dot{\psi}_i\rangle$ with $h_{ji} = \langle\psi_j|\hat{H}|\psi_i\rangle$, and λ is the Lagrange multiplier associated to the holonomic constraint of particle conservation. The derivative of the trace is evaluated using the expression [A.16]. For any given couple of indices k, j we obtain

$$\delta^{(kj)}\mathrm{T}r\ (\dot{\mathrm{f}}\dot{\xi}) = 2\sum_{i} \mathcal{F}_{ijk}\dot{\xi}_{ik}\dot{\xi}_{ji}\delta\xi_{kj} \tag{A.19}$$

The total time derivative in the EL equations [A.17] is derived from eq. [A.16] observing that:

$$\delta\xi_{ij} = \dot{\xi}_{ij}\delta t, \qquad \delta\dot{\xi}_{ij} = \ddot{\xi}_{ij}\delta t \tag{A.20}$$

We then finally obtain the equations of motion presented in Chapter 3.

Appendix B

Computational details

We want here to give a brief overview about the practical use of the algorythms presented in Chapter 3, by giving some hints about the various phases of a typical production calculation.

The very first step involves the set up of the Bravais lattice (i.e. the unit cell and the atomic basis), eventual space symmetries and the k-point grid. The cut-off energy is then fixed according to the “hardness” of the atomic pseudopotentials, and the FFT mesh is generated in real and reciprocal space.

The three dimensions of the mesh are the minimal ones that can fit the cutoff sphere (in reciprocal space) where the charge density is represented, which has typically twice the radius of the cutoff spheres for the wavefunctions. The “big” sphere for the charge density is centered in $\mathbf{G} = 0$, while the “small” spheres are different for each k-point, and are centered in $\mathbf{G} = -\mathbf{k}$. In fact, the code stores the periodic part $u_{n\mathbf{k}}(\mathbf{r})$ of the Bloch state, and this choice ensures that the full wavefunction $\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{kr}}u_{n\mathbf{k}}(\mathbf{r})$ has the correct plane wave cutoff (and thus the total energy has a mathematically exact scaling property). Once the big and small cutoff spheres are defined, time-reversal symmetries are used when possible to reduce the number of independent plane waves that need to be calculated and stored; in particular, for the big sphere, and for k points in the form $\mathbf{k} = (a, b, c)$ where a, b, c are all equal to 0 or $+-1/2$ only one half of the sphere is needed.

Then the G vectors are distributed among the processors, with the following criteria:

- (a) each processor must have an equal amount of G vectors for each k point;
- (b) a single column of the sphere must be on the same processor;
- (c) when time-reversal symmetry is used, columns should be coupled two-by-two.

(a) is mandatory for efficient load-balancing, since the *slowest* PE determines the overall performance. (b) is useful in order to avoid interprocessor communications when the mesh is prepared for 3D FFT (only non-zero plane wave coefficients are stored in memory, and before backward FFT-ing they must be accommodated into the full grid).

Point (c) is a subtle one. The program has only a *complex-to-complex* FFT, while wavefunctions at particular k-points with time-reversal symmetry are *real.* To use a complex-to-complex FFT for transforming real data would be clearly a waste, but LAUTREC uses a trick to avoid performance losses, i.e. for transforming *real* wavefunctions it couples them two-by-two as $\tilde{\psi} = \psi_1 + i\psi_2$, does the FFT, and then recovers the real data as

$$\psi_1 = \frac{1}{2}(\tilde{\psi} + \tilde{\psi}^*), \qquad \psi_2 = \frac{1}{2i}(\tilde{\psi} - \tilde{\psi}^*). \tag{B.1}$$

Now, for the operation to be efficient, $\tilde{\psi}(\mathbf{G})$ and $\tilde{\psi}^*(\mathbf{G})$ need to be on the same processor, and that imposes condition (c).

The amount of electronic states that need to be used is easily evaluated in standard CP dynamics as $N_{band} = N_{el}/2$. For metallic systems, instead, it is costumary to choose a larger value for N_{band}, to ensure that the Fermi-Dirac distribution function is vanishing ($\sim 10^{-8}$) at the upper end of the spectrum, and thus obtain optimal convergence to the electronic gruond state within the algorythm analyzed in Chapter 3. How many unoccupied states should be included? This depends on the system; for aluminum we use $N_{band} \sim 3N_{Al}$, which is a rather large value, because of the high density of states of metallic Al at the Fermi level. For Ag (with 11 electrons per pseudoatom) we find that $N_{band} \sim 6.5N_{Ag}$ is sufficient. For mixed metallic/insulating systems, a good criterion is to take the states needed for the metal plus the bare valence states of the insulator (e.g. $N_{band} \sim 3N_{Al} + 2N_C$ for C_{60}/Al(111)), since the insulator does not contribute significantly to the density of states in the energy range above its valence orbitals.

To initialize the electronic structure we need then a set of trial wavefunctions and occupation numbers. In LAUTREC these can be (i) generated randomly with equal occupations, or (ii) set up from atomic orbitals (which are readily available in pseudopotential data files), with occupation numbers corresponding either to the ground state of isolated atoms or to appropriate ionic configurations.

In (i) one starts in general with an electronic configuration which is very far from the ground state, and should be a bit conservative with the parameters in order to avoid charge-sloshing instabilities [35] in large supercells. We start usually with 10-20 steps of "steepest-descent" dynamics with small timestep (in the range 1-3 a.u. for typical fake masses μ=300 a.u., Q=1000), large electronic smearing ($\sim$ 20 eV), and Gram-Schmidt orthogonalization for the wavefunctions. Then we switch to second-order dynamics with Lagrange multipliers, mass preconditioning and a damping term in both ψ and ξ dynamics:

$$\begin{aligned} \frac{f_{ii}}{2}\mu|\ddot{\psi}_i\rangle &= -f_{ii}\hat{H}_{KS}|\psi_i\rangle - f_{ii}\gamma_\psi\mu|\dot{\psi}_i\rangle + \sum_l \Lambda_{li}|\psi_l\rangle && \text{(B.2)} \\ Q\ddot{\xi}_{ij} &= -(\xi_{ij} - h_{ij} - \lambda) - \gamma_\xi Q\dot{\xi}_{ij} && \text{(B.3)} \end{aligned}$$

We notice that these equations reduce to those of Ref. [35] for standard CP dynamics with f_{ii}=2, and differ by a $\frac{1}{2}$ factor with respect to the original Car and Parrinello paper [33]. By taking a discrete timestep Δ the Verlet equations read then simply:

$$\begin{aligned} \tilde{\psi}^{+}(1+\gamma_{\psi}\Delta) &= 2\psi^{0}-\psi^{-}(1-\gamma_{\psi}\Delta)-\frac{\Delta^{2}}{\mu}(2H_{KS}\psi^{0}) && \text{(B.4)} \\ \tilde{\xi}^{+}(1+\gamma_{\xi}\Delta/2) &= 2\xi^{0}-\xi^{-}(1-\gamma_{\xi}\Delta/2)-\frac{\Delta^{2}}{Q}(\xi_{ij}-h_{ij}) && \text{(B.5)} \end{aligned}$$

Tilded quantities indicate that holonomic constraints need still to be enforced, as explained in Chapter 3. Critical damping is given by the choice:

$$\gamma_{\psi}=\sqrt{\frac{2E_{AG}}{\mu}}, \qquad \gamma_{\xi}=\frac{2}{\sqrt{Q}}, \tag{B.6}$$

where E_{AG} is the "effective gap" discussed in Chapter 3. In practice, γ_{ξ} is always kept constant at the optimal value during minimization to the ground state, while γ_{ψ} is varied, since E_{AG} is well defined in close vicinity of the ground state.

After the first 20 steepest-descent steps (phase 1), wavefunctions and occupancies are still very far from convergence, so typically we perform (phase 2) two damped dynamics runs of 15-20 steps each, with increasing values for Δ (3-5 a.u.), decreasing values for the electronic smearing (5-1 eV) and slightly decreasing γ_{ψ} (in the range 0.04-0.02). Then "production" values can be set for Δ and T (phase 3), with a γ_{ψ} that should be tuned according to the overall convergence rate (some experience is needed in order to ensure optimal damping, but in general 40-80 iterations are sufficient to converge electronic free energy within 10^{-8} Ha of the exact ground state).

If atomic orbitals are used as starting point (option (ii) above), the total energy at startup is already within few percents of the ground state and this allows for some economy in phase 1 and 2. In fact, the first 20 steps (phase 1) can be performed with a Fermi-Dirac smearing which is already rather small (about 1 eV), and phase 2 can then be avoided, thus obtaining a gain of 30-40 steps. In any case, to ensure stability of the wavefunction orthogonalization, vanishing values for f_{ii} are not recommended, and for that reason we impose a lower cutoff of 10^{-8} on the occupation numbers.

We show in Fig. B.1 some examples of the convergence of total energy during phase 3 for highly inhomogeneous metallic systems with large supercells. We report in Table B.1 the computational parameters that have been used in each case. We notice that we did not perform a thorough optimization on these values, but we chose them in a rather conservative way, so that further improvements over the convergence behaviour are well possible. What is interesting about these examples is the stability of the method, which converges smoothly and regularly even in extremely complex systems.

Once the electronic structure has reached the ground state, one can directly relax the atomic positions simultaneously with the electronic and occupational

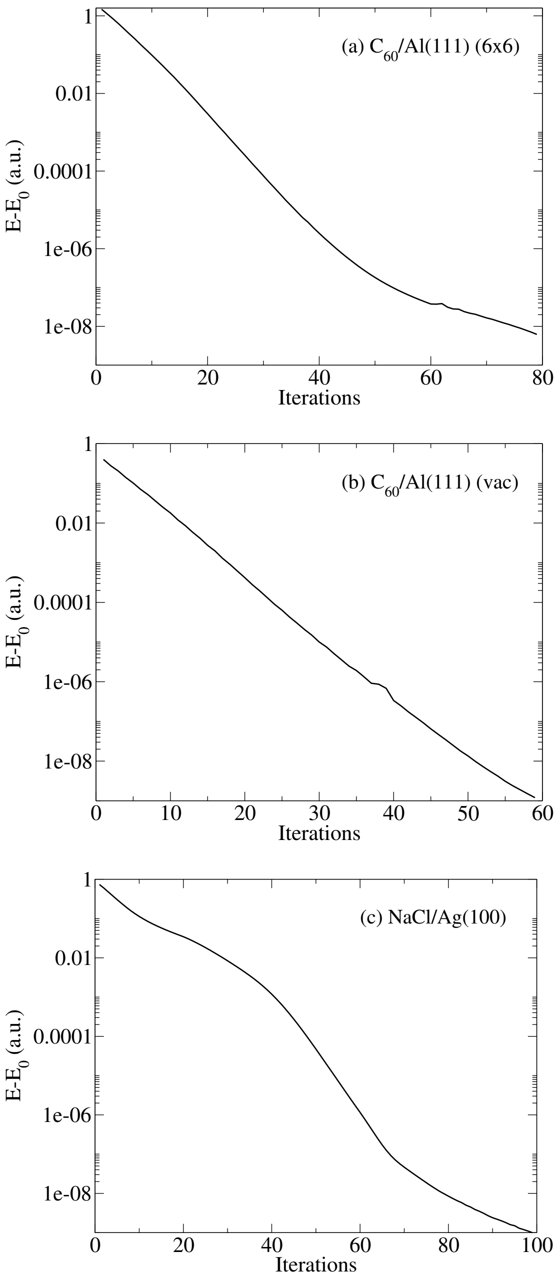


Figure B.1: Convergence of the total energy to the ground state at fixed ions (see Table B.1 for more details).

System	N_{at}	Ω	N_{el}	N_{band}	E_{cut}	E_p	Δ
(a)	60 C + 59 Al	14048.1	417	320	35.0	4.0	7.0
(b)	180 C + 180 Al	42144.3	1260	896	35.0	4.0	7.0
(c)	60 Ag + 1 NaCl	16148.7	668	384	50.0	8.0	6.0

Table B.1: Summary of the computational parameters for three selected systems: (a) C_{60}/Al(111) in the (6×6) supercell (Γ point sampling), (b) C_{60} on a surface vacancy (3 k points), (c) a NaCl molecule on a stepped Ag(100) surface (Γ only). Ω is the supercell volume ($Bohr^3$), other symbols are defined in the text (E_{cut} and E_p are in Rydberg, Δ in a.u.). In all these examples we used $\mu = 300$ and $Q = 1000$, and damping coefficients $\gamma_\psi \sim 0.015 - 0.03$, $\gamma_\xi = 0.06$. The computational load for one iteration at fixed ions was: (a) 2.0 min. (16 T1 processors), (b) 3.3 min. (32 T1 processors), (c) 51 sec. (16 O3K proc.).

degrees of freedom, by choosing some small values $\sim 10^{-3}$ for the friction coefficients.

Appendix C

Performance issues

Even the best electronic structure program is of no practical use unless installed on an efficient parallel server. Nowadays there are two main types of parallel computers, (i) those designed as a whole with proprietary hardware and software (e.g. Cray/SGI, IBM, NEC) and (ii) homogeneous or heterogeneous clusters of workstations connected by vendor-independent networks. The option (ii) is getting more and more popular, essentially because of the lower costs (sometimes an order of magnitude for the same performance) and the high versatility. In addition, cluster computing offers the new opportunity to tailor computer architecture to the specific needs of a given application.

Each computer program performs hardware-intensive operations, but the relative load imposed on different parts of the hardware can vary greatly. There indeed exist expensive servers that provide state-of-the-art performance in every single part of their achitecture, but since the resources are often limited one should obviously invest on the components which are most relevant to the planned activity. Therefore, it is of primary importance to evaluate exactly what are the hardware needs of a given application, and here we will analyze in some detail the performance issues of the LAUTREC program.

LAUTREC is parallelized over the **G** vectors, i.e. the plane wave coefficients for each wavefunction and the FFT mesh are shared among all the processing elements (PE). This means that the scalar products are computed essentially *locally*, while the 3D FFT involves a nonnegligible communication load for transposing of the grid. Let's analyze the workload (CPU and communications) of the 3D FFT routine for the backward FFT of a single wavefunction. We will assume a cubic cell, with n the integer dimension of the mesh, and $V = n^3$ the total number of mesh points. First the plane wave coefficients are accommodated in the reciprocal space grid, and a 1D FFT is performed along all non-zero columns. The workload for a single 1D FFT is $O_{1D} = \alpha n \log_2(n)$, and the non-zero columns are those within a circle of radius $n/4$, so the number of operations in this phase is:

$$O_x = \frac{\pi}{16} n^2 \alpha n \log_2(n) \qquad \text{(C.1)}$$

At this stage, the number of nonzero points are those contained in a cylinder of volume $V' = \frac{\pi}{16}V$, and they need to be transposed into y columns. To do this, each processor has to send across the network $V'(P-1)/P^2$ complex values, and receive the same amount of data. Thus, the global communication workload is

$$C = 2\frac{\pi}{16}V\frac{P-1}{P} \tag{C.2}$$

8-byte words transmitted (one complex is two words). Then the 1D FFT along y direction have to be computed. This involves non-zero y columns which fill up exactly *half* of the grid (the above cylinder with $r_{base} = n/4$ view from one side), giving a CPU workload

$$O_y = \frac{1}{2}n^2\alpha n\log_2(n) \tag{C.3}$$

For simplicity we assume here that n is an integer multiple of P. In this case the layout in y direction is such that the transposition to the final z leading direction is *local* on separate processors, without communications. Then to finalize the 3D FFT one has to compute the 1D FFT along z, which is performed on *all* columns (there are no more zero columns at this point):

$$O_z = n^2\alpha n\log_2(n) \tag{C.4}$$

The *total* CPU workload is given by

$$O_{3D} = O_x + O_y + O_z = \frac{\alpha}{3}(\frac{\pi}{16} + \frac{1}{2} + 1)V\log_2(V) \tag{C.5}$$

We can then define γ_a as the ratio between CPU operations and words communicated through the network:

$$\gamma_a^{FFT} \sim \frac{0.56\alpha V\log_2(V)}{0.40V(P-1)/P} = 1.44\alpha\log_2(V)\frac{P}{P-1} \tag{C.6}$$

In typical production calculations $\log_2(V) \sim 20$, and 16-32 processors are used so we can consider $(P-1)/P \sim 1$; finally, for a complex-to-complex FFT $\alpha = 5$, and thus γ_a^{FFT} has a value around 140-150. This means that, during 3D FFTs, for every 150 floating-point operations one single processing element must send and receive an 8-byte word.

Unfortunately, there is no simple way to superimpose communications and operations, so during interprocessor data transfer the CPUs remain idle. This is the reason why the *network* performance is often the most serious bottleneck of LAUTREC, but also of any other electronic structure code which adopts the same parallelization strategy. It is easy to see that standard 100 Mbit/s (1.5 Mword/s) Ethernet connections are poorly suited: nowadays CPUs can sustain more than 1 Gflop/s (FLoating point OPerations), and the time necessary to communicate one single word is then almost five times longer than that necessary to perform 140 operations!

A parameter which is relevant for evaluating the suitability of a given architecture to a chosen algorythm implemented in LAUTREC is γ_m, defined as the CPU performance in Mflop/s divided by the network performance in Mword/s. Then one should ensure that $\gamma_m < \gamma_a$ in order to avoid a substantial waste of CPU power. For the above example $\gamma_m = 640$, which is much larger than γ_a^{FFT}, i.e. high-performance CPUs connected with Ethernet are not a wise choice for LAUTREC. For the study of the C_{60}/Al(111) system we used extensively the T1 cluster, which provided a much better γ_m (of about 125), and thus allowed for fairly good performance. One should notice also the 3D FFTs are only one part of the total workload, and indeed the most communication-intensive one. For large supercells, like the C_{60}/Al(111) (6×6) one, the CPU workload for performing scalar products (matrix-matrix multiplications) largely dominated over FFTs (the global wall clock timing for the FFTs *including* communications was typically three times shorter than matrix-matrix operations), and the overall communication overhead was estimated to be $\sim 10-15$ % only.

Bibliography

[1] M. Stengel and A. De Vita, Phys. Rev. B **62**, 15283 (2000).

[2] S. Schintke, S. Messerli, M. Pivetta, F. Patthey, L. Libioulle, M. Stengel, A. De Vita and W. D. Schneider, Phys. Rev. Lett. **87**, 276801 (2001).

[3] M. Vladimirova, M. Stengel, A. De Vita, A. Baldereschi, M. Böhringer, K. Morgenstern, R. Berndt and W.-D. Schneider, Europhys. Lett. **56**, 254 (2001).

[4] R. Gruber, P. Volgers, A. De Vita, M. Stengel, T.-M. Tran, Fut. Gen. Comp. Systems **19**, 111 (2003).

[5] M. Stengel, A. De Vita and A. Baldereschi, Phys. Rev. Lett. **91**, 166101 (2003).

[6] R. E. Smalley and B. I. Yakobson, Solid State Comm. **107**, 597 (1998).

[7] R. M. Metzger, J. of Solid State Chem. **168**, 696 (2002).

[8] S.-W. Hla and K.-H. Rieder, Superlattices and Microstructures **31**, 63 (2002). *Available online at http://www.idealibrary.com.*

[9] R. Stadler and M. Forshaw, Physica E **13**, 930 (2002).

[10] W.-D. Schneider, Surf. Sci. **514**, 74 (2002).

[11] J. V. Barth, J. Weckesser, G. Trimarchi, M. Vladimirova, A. De Vita, C. Z. Cai, H. Brune, P. Gunter and K. Kern, J. Am. Chem. Soc. **124**, 7991 (2002).

[12] L. C. Ciacchi, W. Pompe and A. De Vita, J. Phys. Chem B **107**, 1755 (2003).

[13] S. Dieckhoff *et al.*, Surf. Sci. **279**, 233 (1992).

[14] D. Ochs *et al.*, Surf. Sci. **383**, 162 (1997).

[15] M. Kiguchi *et al.*, Surf. Sci. **522**, 84 (2003).

[16] P.Hohenberg and W.Kohn, Phys. Rev. **136**, B864 (1964).

[17] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[18] L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).

[19] N. D. Mermin, Phys. Rev. **137**, A1441 (1965).

[20] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[21] J. Perdew and A.Zunger, Phys. Rev. B **23**, 5048 (1981).

[22] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).

[23] G. Onida, L. Reining and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).

[24] P. García González and R. W. Godby, Comp. Phys. Comm. **137**, 108 (2001).

[25] J. P. Perdew and M. Levy, Phys. Rev. Lett. **51**, 1884 (1983).

[26] R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).

[27] V. Fiorentini and A. Baldereschi, J. Phys.: Condens. Matter **4**, 5967 (1992).

[28] S. G. Louie, S. Froyen and M. L. Cohen, Phys. Rev. B **26**, 1738 (1982).

[29] D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990).

[30] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).

[31] P. E. Blöchl e M. Parrinello, Phys. Rev. B **45**, 9413 (1992).

[32] S. Nosé, Molec. Phys. **52**, 255 (1984).

[33] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).

[34] G. Pastore, E. Smargiassi and F. Buda, Phys. Rev. A **44**, 6334 (1991).

[35] F. Tassone, F. Mauri and R. Car, Phys. Rev. B **50**, 10561 (1994).

[36] N. Marzari, D. Vanderbilt and M. C. Payne, Phys. Rev. Lett. **79**, 1337 (1997).

[37] M. P. Grumbach *et al.*, J. Phys.: Condens. Matter **6**, 1999 (1994).

[38] J. VandeVondele and A. De Vita, Phys. Rev. B **60**, 13241 (1999).

[39] K. Hedberg *et al.*, Science **254**, 410 (1991).

[40] J. E. Fischer *et al.*, Science **252**, 1288 (1991).

[41] C. Pan *et al.*, J. Phys. Chem. **95**, 2944 (1991).

[42] N. Troullier and J. L. Martins, Phys. Rev. B **46**, 1754 (1992).

[43] M. Hasegawa *et al.*, J. Chem. Phys. **119**, 1386 (2003).

[44] C. Joachim *et al.*, Nature **408**, 541 (2000).

[45] O. Gunnarsson, Rev. Mod. Phys. **69**, 575 (1997).

[46] A.J.Maxwell *et al.*, Phys. Rev. B **57**, 7312 (1998).

[47] M. Pedio *et al.*, Phys. Rev. Lett. **85**, 1040 (2000).

[48] P.W. Murray *et al.*, Phys. Rev. B **55**, 9360 (1997).

[49] J. Weckesser *et al.*, Phys. Rev. B **64**, 161403 (2001).

[50] C. Rogero *et al.*, J. of Chem. Phys. **116**, 832 (2002).

[51] M.K.-J.Johansson *et al.*, Surf. Sci. **397**, 314 (1998).

[52] A.J.Maxwell *et al.*, Phys. Rev. B **52**, R5546 (1995).

[53] M.K.-J.Johansson *et al.*, Phys. Rev. B **54**, 13472 (1996).

[54] R.Fasel *et al.*, Phys. Rev. Lett. **76**, 4733 (1996).

[55] R.Stumpf and M.Scheffler, Phys. Rev. B **53**, 4958 (1996).

[56] M.Fuchs and M.Scheffler, Comp. Phys. Comm. **119**, 67 (1999).

[57] A.P.Seitsonen *et al.*, J. Chem. Phys. **103**, 8075 (1995).

[58] B. H. Mahan, R. J. Myers, *University Chemistry*, Benjamin/Commings Publishing (1987).

[59] M.T.Cuberes *et al.*, Appl. Phys. A **66**, S669 (1998).

[60] E.M. King *et al.*, J. Phys. Chem. B **105**, 641 (2001).

[61] R. Duschek *et al.*, Chem. Phys. Lett. **318**, 43 (2000).

[62] R.I.R. Blyth *et al.*, J. Chem. Phys. **114**, 935 (2001).

[63] S. Mezhenny *et al.*, J. Am. Chem. Soc. **124**, 14202 (2002).

[64] S. Altieri *et al.*, Phys. Rev. B **59**, R2517 (1999).

[65] S. Valeri *et al.*, Surf. Sci. **507**, 311 (2002).

[66] S. Valeri *et al.*, Phys. Rev. B **65**, 245410 (2002).

[67] L. Giordano *et al.*, Phys. Rev. B **67**, 045410 (2003).

[68] S. Altieri, L. H. Tjeng and G. A. Sawatzky, Phys. Rev. B **61**, 16948 (2000).

[69] J. Wollschläger *et al.*, Appl. Surf. Sci. **142**, 129 (1999).

[70] M. Kiguchi *et al.*, Surf. Sci. **512**, 97 (2002).

[71] M. Sgroi, C. Pisani and M. Busso, Thin Solid Films **400**, 64 (2001).

[72] C. Giovanardi *et al.*, Surf. Sci. **505**, L209 (2002).

[73] C. Li, R. Wu and A. J. Freeman, Phys. Rev. B **48**, 8317 (1993).

[74] C. Noguera, Surf. Rev. and Lett. **8**, 121 (2001).

[75] U. Schönberger and F. Aryasetiawan, Phys. Rev. B **52**, 8788 (1995).

[76] S. Pugh and M. J. Gillan, Surf. Sci. **320**, 331 (1994).

[77] G. Cappellini *et al.*, J. Phys.: Condens. Matter **12**, 3671 (2000).

[78] P. K. de Boer and R. A. de Groot, J. Phys.: Condens. Matter **10**, 10241 (1998).

[79] P. V. Sushko, A. L. Schluger and C. R. A. Catlow, Surf. Sci. **450**, 153 (2000).

[80] C. Satoko *et al.*, J. Phys. Soc. Japan **45**, 1333 (1978).

[81] L. N. Kantorovich *et al.*, Faraday Discuss. **114**, 173 (1999).

[82] M. Klaua *et al.*, Phys. Rev. B **64**, 134411 (2001).

[83] O. Robach *et al.*, Surf. Sci. **401**, 227 (1998).

[84] P. A. Cox and A. A. Williams, Surf. Sci. **175**, L782 (1986).

[85] V. E. Henrich, G. Dresselhaus and H.J. Zeiger, Phys. Rev. B **22**, 4764 (1980).

[86] H. Erschbaumer *et al.*, Surf. Sci. **243**, 317 (1991).

[87] C. Kittel, *Introduction to Solid State Physics*, John Wiley & Sons, Inc. (1986).

[88] J.D. Jackson, *Classical Electrodynamics*.

[89] W.H. Press, B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, *Numerical recipes*, Cambridge University Press (1986).

[90] N. W. Ashcroft, N. D. Mermin, *Solid State Physics*, Saunders College Publishing (1976).

[91] A. Zangwill, *Physics at Surfaces*, Cambridge University Press (1988).

[92] M. Scheffler e C. Stampfl in *Handbook of Surface Science*, ed. K. Horn e M. Scheffler, Elsevier, Amsterdam (1999).

[93] M. W. Cole and M. H. Cohen, Phys. Rev. Lett **23**, 1238 (1969).

[94] E. V. Chulkov, V. M. Silkin and P. M. Echenique, Surf. Sci. **391**, L1217 (1997).

[95] E. V. Chulkov, V. M. Silkin and P. M. Echenique, Surf. Sci. **437**, 330 (1999).

[96] V. Dose *et al.*, Phys. Rev. Lett. **52**, 1919 (1984).

[97] B. Reihl, K. H. Frank and R. R. Schlitter, Phys. Rev. B **30**, 7328 (1984).

[98] P. A. Serena, J. M. Soler and N. García, Phys. Rev. B **34**, 6767 (1986).

[99] A. G. Eguiluz, Phys. Rev. Lett. **68**, 1359 (1992).

[100] I. D. White *et al.*, Phys. Rev. Lett. **80**, 4265 (1998).

[101] M. Weinert, S. L. Hulbert and P. D. Johnson, Phys. Rev. Lett **55**, 2055 (1985).

[102] Z. Li and S. Gao, Phys. Rev. B **50**, 15349 (1994).

[103] S. Jørgensen *et al.*, Chem. Phys. **278**, 53 (2002).

[104] D. C. Marinica *et al.*, Phys. Rev. Lett. **89**, 046802 (2002); D. C. Marinica *et al.*, Surf. Sci. **540**, 457 (2003).

[105] D. Straub and F. J. Himpsel, Phys. Rev. Lett. **52**, 1922 (1984).

[106] G. Binnig *et al.*, Phys. Rev. Lett. **55**, 991 (1985).

[107] J. Tersoff, Phys. Rev. Lett. **52**, 465 (1984).

[108] R. Hesper, L. H. Tjeng and G. A. Sawatzky, Europhys. Lett. **40**, 177 (1997).

[109] Y. Murata *et al.*, J. Phys. Soc. Japan **70**, 793 (2001).

[110] J. Tersoff and D. R. Hamann, Phys. Rev. B **31**, 805 (1985).

[111] W. Hebenstreit *et al.*, Surf. Sci. **424**, L321 (1999).

[112] N.-P. Wang *et al.*, Phys. Rev. B **67**, 115111 (2003).

[113] X. Gonze, J.-M. Beuken, R. Caracas, F. Detraux, M. Fuchs, G.-M. Rignanese, L. Sindic, M. Verstraete, G. Zerah, F. Jollet, M. Torrent, A. Roy, M. Mikami, Ph. Ghosez, J.-Y. Raty, D.C. Allan., Computational Materials Science **25**, 478-492 (2002).

[114] M. Kiguchi *et al.*, Phys. Rev. Lett. **90**, 196803 (2003).

[115] M. Rohlfing *et al.*, Phys. Rev. Lett. **91**, 256802 (2003).

[116] S. Schintke, Ph.D. thesis, Université de Lausanne, IPMC-BSP, CH-1015 Lausanne (2001).

Remerciements

Je remercie tout d'abord le Professeur Baldereschi de m'avoir accueilli dans son groupe, et d'avoir suivi directement une partie importante de mon travail. Son approche scientifique et humaine de la recherche, que j'ai eu la chance d'apprécier au travers de nombreuses et passionnantes discussions, m'a apporté une richesse qui m'est difficile à quantifier.

J'aimerais exprimer toute ma gratitude au Dr. Alessandro De Vita, qui m'a appris le métier difficile du chercheur, et qui m'a toujours soutenu avec enthousiasme et amitié dans mon travail. Cette thèse n'aurait jamais pu avoir lieu sans son aide, sa compétence et sa disponibilité.

Je remercie le Dr. M. Peressi pour les discussions stimulantes que nous avons eues pendant ses visites à l'EPFL, ainsi que le Dr. A. Filippetti, qui m'a mis à disposition son programme et m'a introdui aux problematiques liées à la correlation électronique dans les matériaux.

C'est un plaisir de remercier le Professeur W.-D. Schneider et tout son groupe de recherche pour avoir fourni les résultats qui ont inspiré une partie de cette these.

Je remercie aussi le Professeur M. Payne, le Dr. W. Andreoni et le Prof. H. Brune, qui ont accepté d'être membres du jury de ma thèse.

Je tiens à remercier tous mes collègues et amis de l'IRRMA, qui ont tous contribué d'une manière ou de l'autre à la réussite de ce travail, sous le point de vue professionnel et non professionnel, en particulier le Dr. Paolo Umari, le Dr. Reiner Härle, et le Dr. Masha Vladimirova.

Un merci spécial va enfin à Peggy, à ma mère qui m'a toujours encouragé, et à tous les amis grimpeurs avec qui j'ai eu la chance de partager des bons moments en montagne pendant ces quatre ans.

Curriculum Vitae

Nom et prénom	Stengel Massimiliano
Date de naissance	20 avril 1974
Lieu d'origine	Trieste (Italie)
Nationalite	Italienne
État civil	Célibataire

2000-2004	Assistant-doctorant à l'*Institut Romand de Recherche Numérique en Physique des Matériaux* (IRRMA), EPFL.
2001-2002	Assistant pour les cours "Physique I" du Prof. H. Brune et "Experimentations Numeriques" du Prof. A. Baldereschi.
December 1999	Laurea in Fisica (Diplôme en Physique) *summa cum laude* à l'Università di Trieste (Italie).
1993-1999	Études de physique, Università di Trieste (Italie).